DIAGNOSTIC MEASUREMENT FROM A STANDARDIZED MATH ACHIEVEMENT TEST USING MULTIDIMENSIONAL LATENT TRAIT MODELS

A Thesis
Presented to
The Academic Faculty

by

HeaWon Jun

In Partial Fulfillment
of the Requirements for the Degree
Master of Philosophy in the
School of Psychology

Georgia Institute of Technology
May 2014

DIAGNOSTIC MEASUREMENT FROM A STANDARDIZED MATH ACHIEVEMENT TEST USING MULTIDIMENSIONAL LATENT TRAIT MODELS

Approved by:

Dr. Susan Embretson, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Larry James
School of Psychology
Georgia Institute of Technology

Dr. Christopher Hertzog
School of Psychology
Georgia Institute of Technology

Date Approved: April 18, 2013

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Susan Embretson for her support, guidance, and encouragement throughout my research. I would also like to thank Dr. Larry James and Dr. Chris Hertzog for their guidance and persistent help for thesis work. Without your help, this paper would not have been possible. Special thanks go to my family and friends who supported me in writing, and incented me to strive towards my goal.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
SUMMARY	viii
CHAPTER1: INTRODUCTION	1
CHAPTER2: THEORETICAL BACKGROUND	5
Multidimensional Item Response Theory Models	5
Compensatory MIRT Models	5
Noncompensatory MIRT Models	7
The DINA Model	9
CHAPTER3: METHOD	11
Participants	11
Testing Materials	11
Procedure	13
Testing models	13
CHAPTER4: RESULTS	14
Descriptive Statistics	14
The EMIRT Model	15
Unidimensional IRT Model	15
Full Information Bifactor Model	17
The CMIRT Model	19
Diagnosis from the CMIRT Model	25

Dimension and pattern probabilities 25

The DINA model 30

Comparison of Dimension Patterns from both the DINA and CMIRT models 31

The MLTM-D 41

Diagnosis from the MLTM-D 43

Comparison of Attribute Thetas from the MLTM-D and the CMIRT model 45

CHAPTER5: DISCUSSION 46

APPENDIX A: MODEL PARAMETER ESTIMATES AND FIT INDICES 48

REFERENCES 97

LIST OF TABLES

Page

Table 1: Goodness of fit comparison among IRT models 16

Table 2: Descriptive statistics of theta estimates for four components from augmented blueprint based CMIRT model 22

Table 3: Correlation coefficients among theta estimates from augmented blueprint based CMIRT model 24

Table 4: Descriptive statistics of dimension probabilities from augmented blueprint based CMIRT model 27

Table 5: Distribution in dimension patterns at different dimension proficiency cutlines: augmented blueprint-based CMIRT and DINA 29

Table 6: CMIRT attribute probabilities using 58% cutline 34

Table 7: CMIRT attribute probabilities using 63% cutline 35

Table 8: CMIRT attribute probabilities using 68% cutline 36

Table 9: CMIRT attribute probabilities using 70% cutline 37

Table 10: CMIRT attribute probabilities using 73% cutline 38

Table 11: CMIRT attribute probabilities using 78% cutline 39

Table 12: Phi coefficient coefficients between attribute patterns from the DINA and CMIRT models 40

Table 13: Descriptive statistics for MLTM-D component item estimates 41

Table 14: Descriptive statistics of theta estimates for four components from MLTM-D 42

Table 15: Correlation coefficients among theta estimates from MLTM-D 42

LIST OF FIGURES

Page

Figure 1: Hierarchical structure of standards, benchmarks, and indicators on a math test 12

Figure 2: Configuration of Bifactor model 18

Figure 3: Configuration of CMIRT model based on the original test blueprint 20

Figure 4: Configuration of CMIRT model based on the augmented test blueprint 21

Figure 5: Distribution graphs of the theta estimates from the augmented blueprint based CMIRT model 23

Figure 6: Patterns of CMIRT theta estimates on four components from students with same total score 25

Figure 7: Patterns of CMIRT dimension probabilities of four student samples 28

Figure 8: Regression of observed proportion passing each item on MLTM-D prediction across categories 44

Figure 9: MLTM-D component competency patterns of four student samples 45

SUMMARY

The present study compares applications of continuous multidimensional item response theory (MIRT) models for their diagnostic potential. Typically, MIRT models have not been used for diagnosing the possession of skills or attributes by students, but several researchers have suggested that they can potentially be used for this purpose (e.g., Stout, 2007; Wainer, Vevea, Camacho, Reeve, Rosa, Nelson, Swygert, & Thissen, 2001). This study applies MIRT models to a standardized eighth grade mathematics achievement test that was constructed based on a hierarchically-structured blueprint consisting of standards, benchmarks, and indicators. Only the highest level, consisting of four standards, was used to define the dimensions. The confirmatory models were defined using the standards that had been scored for involvement in each item. For the current study, the exploratory MIRT (EMIRT) model was interpreted with respect to the dimensions. Then, the compensatory and confirmatory MIRT (CMIRT) models and the full information bifactor model were fitted. The interpretation of dimensions, empirical reliabilities of person estimates, and test- and item-fit were examined. Also, dimension and pattern probabilities were obtained for determining their diagnostic potential. Last, a noncompensatory MIRT model (MLTM-D; Embretson & Yang, 2011) and the DINA model (Haertel, 1989; Junker & Sijtsma, 2001) in use as diagnostic models were analyzed to compare pattern probabilities with the compensatory CMIRT model.

CHAPTER 1

INTRODUCTION

A tool for measuring test takers' processing knowledge, strategies, and skills while solving problems is diagnostic classification models (DCMs). These models have potential to provide valuable information about students' strengths and weaknesses in their specific area of skills or attributes involved in items.

DCMs require support for two aspects of construct validity research. Two major types of evidence for construct validity (Standards for Educational and Psychological Tests, 1999) are the content and response process aspects. Taken together, these two aspects define construct representation (Embretson, 1983) which is defined as the processes, strategies and knowledge that the examinees apply to solve the items. Meanwhile, the attainment of construct validity especially for achievement tests depends on the appropriateness and completeness of construct representation (Yang and Embretson, 2007). The appropriateness of construct representation indicates suitableness of form or symbol system which is employed to describe a construct and the completeness of construct representation is a characteristic obtained by identifying all relevant constructs that a test measures and limiting the impact of irrelevant constructs. DiBello, Stout, and Roussos (1995) claimed that completeness of construct representation is achieved by identifying specific knowledge or strategies consisting of a task, not in the task itself, and analyzing data with the specific task information using appropriate psychometric models. Embretson (1983) also stated that searching for components and their theoretical mechanisms that consist of test performance is an important step to identify construct validity.

Diagnostic models provide person estimates on the possession of attributes, skills, or even strategies by relating item response probabilities to skills or strategies that are req uired to solve items. Diagnostic models can be classified into two different types: diagnos tic latent classification models and diagnostic latent trait models. Most diagnostic models are latent classification models; these models are probabilistic in that they directly link ob servable item responses and an unobservable latent variable (i.e., the latent class). The mo dels are confirmatory in that latent variables an item involves are known in advance and specified for model estimation (Rupp, Templin, & Henson, 2010). The diagnostic latent c lassification models assume categorical latent variables instead of continuous latent traits, providing classification information with respect to various latent characteristics. On the other hand, the diagnostic latent trait models offer continuous trait level estimates in com ponents, attributes, or skills rather than classification information. Diagnostic latent trait models also estimate examinees' global competency levels on dimensions rather than a bi nary mastery level of the attributes or skills related to dimensions. However, more specific skills can be estimated if the skills are linearly ordered on the latent dimensions.

The reparameterized unified model (RUM: Hartz, 2002), the reduced RUM (Hartz , Roussos, & Stout, 2002), the general diagnostic model (GDM; von Davier & Yamamoto , 2007; von Davier, 2005), the deterministic inputs, noisy "and" gate (DINA; Haertel, 198 9; Junker & Sijtsma, 2001), and the log-linear cognitive diagnosis model (LCDM: Henso n, Templin, & Willse, 2009) are all examples of diagnostic classification models. On the other hand, the linear logistic test model (LLTM; Fisher, 1973), the multidimensional logi stic model (Reckase &McKinley, 1982), the multicomponent Rasch model (Stegelmann, 1983), multicomponent latent trait model (MLTM: Whitely, 1980), general component lat ent trait model (GLTM; Embretson, 1984), and multicomponent latent trait model for diagnosis (MLTM-D: Embretson & Yang, 2011) are categorized as diagnostic latent trait models because they have diagnostic potential from latent dimensions.

Because most diagnostic models provide estimates of attributes or skills possessed by examinees, the models can give students or instructors relevant remedial instruction information on the skills involved in the test. For this reason, diagnostic classification models have been actively studied and further applied to several tests (Pellegrino, 2002; Gierl & Zhou, 2008); multidimensional latent trait models, allowing only global dimension competency level to be estimated, were believed to have only limited use for diagnostic purposes. However, other researchers focused on the potentially diagnostic use of latent trait models under certain conditions (e.g., Stout, 2007; Wainer et al., 2001). For example, Wainer et al. (2001) noted that MIRT models can offer available diagnostic information because person scores for subcomponents can be obtained, and Stout (2007) stated that, given confirmatory knowledge, MIRT models have potential for diagnosis by observing and interpreting how dimensions are combined. Embretson and Yang (2011) pointed out that the MIRT models are more parsimonious than most classification models because they have fewer parameters to be estimated. Although diagnostic classification models have advantages for diagnosis, their application to high-stakes achievement tests may be limited when too many attributes are involved in the items. For example, if 20 or more attributes were involved in a test, the possible maximum number of attribute patterns for each binary item would be 2^{20}, which is very large. Too many attribute patterns could decrease the reliability of the parameter estimates. Compared to the classification models, the MIRT models only obtain trait levels of components as dimensions and then interpret these trait estimates themselves or the estimates through relationships of components and skills or attributes.

The purpose of the current study is to apply several different MIRT models to a standardized mathematics achievement test and to examine their diagnostic potential. A brief overview of diagnostic latent trait models precedes the explanation of several testing models that will be analyzed in this paper. The unidimensional IRT models, EMIRT and CMIRT models, and the MLTM-D will be fitted to the data; the interpretability of the dimensions, empirical reliabilities of theta estimates, and fit indices will be compared across models. Then, the dimension and pattern probabilities will both be computed, which will then be compared to those from the DINA model and the MLTM-D to test the similarity of the attribute patterns.

CHAPTER 2

THEORETICAL BACKGROUND

Multidimensional Item Response Theory Models

Many large-scale tests include items that involve multiple skills or attributes. If examinees and items differ in those skills or attributes, MIRT models may be considered for estimating person abilities (Embretson, 2000). MIRT models include two or more item discrimination parameters and trait scores; the item discrimination parameters in the models represent the relative impact of the dimensions on items, and trait scores are locations on the dimensions for each person. The multidimensional latent trait models include both compensatory and noncompensatory MIRT models. In compensatory models, a high ability on a component can compensate for low abilities on other components required by an item. In noncompensatory models, some minimum ability on each of the relevant components is required (Embretson, 2000). Compensatory MIRT models, such as multidimensional logistic models and the multidimensional normal ogive model, and noncompensatory MIRT models such as the MLTM (Embretson, 1980), the GLTM (Embretson, 1984), and the MLTM-D (Embretson & Yang, 2011) will be described in the following sections, along with one of the most straightforward cognitive diagnostic classification models, the DINA model.

Compensatory MIRT Models

An example of compensatory models, the two-parameter logistic (2PL) compensatory MIRT model, is given as:

$$P(X_{is}=1|\underline{\theta}_s,\delta_i,\underline{\alpha}_i)=\frac{\exp(\sum_m \alpha_{im}\theta_{sm}+\delta_i)}{1+\exp(\sum \alpha_{im}\theta_{sm}+\delta_i)}\ ,$$

where δ_i is the easiness intercept, θ_{sm} is the trait score of examinee *s* on dimension *m*, and α_{im} is the weight of dimension *m* on item *i*. In this model, the probability that a person responds to an item successfully depends on the item difficulty and a weighted combination of the abilities (Reckase & McKinley, 1991). Specifically, the higher the weight on one dimension relative to another dimension, the more influential that trait is on the item; at the same time, as a person possesses more ability on a given dimension, the probability that the person completes an item successfully increases. A three-parameter logistic (3PL) compensatory MIRT model is not very different from the 2PL MIRT model, except for the inclusion of a guessing parameter, r_i, in the model. The 3PL MIRT model is represented as follows:

$$P(X_{IS}=1|\underline{\theta}_s,\delta_i,\underline{\alpha}_i,r_i)=r_i+(1-r_i)\frac{\exp(\sum_m \alpha_{im}\theta_{sm}+\delta_i)}{1+\exp(\sum \alpha_{im}\theta_{sm}+\delta_i)}\ .$$

The multidimensional normal ogive IRT model (Bock, Gibbons, & Muraki, 1988) is similar to the multidimensional logistic models, but its function uses a cumulative normal distribution. Bock et al. (1988) treated the multidimensional normal ogive IRT model as a full information factor analysis model. Compensatory MIRT models can be classified as exploratory or confirmatory models. In exploratory models, the number of dimensions is determined by investigating the fit of different models to the data in the absences of hypothesized theories to specify constraints on the estimation process. In contrast, confirmatory models specify the relationship of the items to the dimensions before parameter estimation.

The current study tests unidimensional IRT, a full information bifactor IRT, EMIRT, and CMIRT models using IRTPRO 2.1 (Thissen & Cai, 2011). For item calibration, IRTPRO has options to use the method of maximum likelihood (ML), maximum *a posteriori* (MAP), or alternative computational methods such as the Bock-Aitkin expectation maximization (BAEM; Bock & Aitkin, 1981), adaptive quadrature expectation maximization (ADQEM; Schilling & Bock, 2005), and Metropolis-Hastings Robbins-Monro (MH-RM; Cai, 2010) estimation methods. The estimation of person scores can be done using maximum *a posteriori* (MAP), or expected *a posteriori* (EAP) estimation.

Noncompensatory MIRT Models

Each of the MLTM (Whitely, 1980), GLTM (Embretson, 1984), and MLTM-D (Embretson & Yang, 2011) is a noncompensatory multidimensional IRT model. The MLTM estimates multiple component difficulties and multiple component trait levels underlying an item. The MLTM is given as follows:

$$P(X_{isT} = 1 | \underline{\theta}_s, \underline{\beta}_i) = \prod_m \frac{\exp(\theta_{sm} - \beta_{im})}{1 + \exp(\theta_{sm} - \beta_{im})} ,$$

where θ_{sm}, and β_{im} represent the trait level of person *s* on the component *m*, and the difficulty of item *i* on component *m*, respectively. The MLTM measures the probability that examinee *s* completes item *i* successfully, X_{isT} , as the product of the probabilities that the examinee succeeds on the components involved. The GLTM is a generalized form of the MLTM, in which the β_{im} term is replaced by $\sum_k \tau_{km} q_{ikm} + q_{0m}$, the weighted sum of underlying stimulus factors (q_{ikm}) to represent scored attributes. Both the MLTM and GLTM require two levels of data - total item response and component responses - to estimate the person and item level component parameters. Although some attempts have been made to investigate whether the MLTM and GLTM without subtask responses are practical, the results were not consistent (Yang & Embretson, 2007; Embretson, 1995; Embretson & McCollam, 2000; Bolt & Lall, 2003).

Another noncompensatory MIRT model, the MLTM-D is a diagnostic model in which two levels of a hierarchical structure, components and attributes, determine a response probability. The MLTM-D is valid when two or more components affect problem solving and when the difficulties of those components, in turn, are influenced by those of lower-level attributes. The C- and Q- matrices are data matrices that show the relationships of the components to specific items and the relationships between the attributes and the components, respectively In these matrices, if an item is considered to be relevant to two components among all the identified components, the corresponding two elements of the matrix are indicated by 1 and the other elements by 0; the attributes are indicated for each component in the same way. The probability of solving an item successfully in MLTM-D is shown as follows:

$$P(X_{ij}=1)=\prod_{m}\left[\frac{1}{1+\exp\{-1.7(\theta_{jm}-\sum_{k}\eta_{mk}q_{imk}+\eta_{0})\}}\right]^{c_{im}},$$

where θ_{jm}, q_{imk}, η_{mk}, and c_{im} indicate the ability of person *j* on component *m*, the score of attribute *k* of component *m* in item *i*, the weight of attribute *k* on component *m*, and the involvement of component *m* in item *i*, respectively. In the MLTM-D, the ability estimates are obtained for each component. By linking the component to the specific attributes relevant to the component and observing the association between the component and attributes, the level of the attribute can be determined. Thus, if a mastery level for each component is obtained, we can locate examinees on the attribute scale, as in the LLTM. This model allows information on the level of the attributes that an examinee possesses without directly estimating score levels on attributes. Because of its hierarchical structure, the MLTM-D requires a component structure that represents the involvement of attribute k. Therefore, the MLTM-D requires **C-** and **Q-** matrices for model identification. For the current study, item and person parameters of the MLTM-D were estimated using marginal maximum likelihood method, and the measurement scale was fixed by setting the mean of the item difficulties and the item discrimination to 0 and 1, respectively. Among two levels of diagnostic information, component and attributes, that MLTM-D provides, only the estimates on component level are used.

The DINA Model

One of the simplest conjunctive classification diagnostic models is the deterministic inputs noisy "and" gate (DINA; Haertel, 1989; Junker and Sijtsma, 2001) model. In the DINA model, an examinee's binary skill vector, denoted by $\alpha_s = \{\alpha_{sk}\}$, is a matrix representing whether the *s*th examinee mastered the *k*th skill required for answering an item correctly. If the *k*th skill is mastered by examinee *s*, the corresponding element in the vector is denoted by 1; otherwise, it is 0. The DINA model requires a specification of the **Q-**matrix representing which skills an item requires for a correct response. If assuming that item *i* requires skill *k* to be solved, the corresponding element of *i* x *k* is replaced by 1; otherwise, it is 0. This matrix also creates a binary matrix consisting of 0 and 1. Given these two matrices, the DINA model produces a latent response vector $\eta_i = \{\eta_{is}\}$, as follows:

$$\eta_{is} = \prod_{k=1}^{K} \alpha_{sk}^{\;\;ik} .$$

If all the required skills for answering item *i* correctly are possessed by examinee *s*, η_{is} will equal 1, and if at least one of the required skills is not possessed by examinee *i*, η_{is} will become 0. Therefore, an examinee who answers an item correctly is expected to have all the required skills for that item. The latent response vector represents an ideal response pattern, but the underlying process of the DINA model is stochastic. The stochastic process is determined by two parameters, slip (s_i) and guessing (g_i) parameters. In the former, an examinee who acquires all the required skills for answering an item may still answer it incorrectly (i.e., slip), and in the latter, an examinee who does not possess at least one among all the required skills may answer an item correctly by guessing. In the DINA model, these two parameters, s_i and g_i , are written as $P(X_{is} = 0 \ | \eta_{is} = 1)$ and $P(X_{is} = 1|$ $\eta_{is} = 1)$, respectively, in which X_{is} represents the item response of examinee *s* to item *i*. Specifically, the probability that examinee *s* with a specific skill pattern, α_s , answers item *i* correctly is defined by $P_i(\alpha_s) = P(X_{is} = 1|\alpha_s) = g_i^{1-\eta_{is}}(1-s_i)^{\eta_{is}}$. If η_{is} is equal to 1, the probability is determined only by the slip parameter, s_i , and if η_{is} is equal to 0, it is determined by the guessing parameter, g_i . The DINA model is a simple and parsimonious model, allowing only two parameters, s_i and g_i , to be estimated for each item even when many attributes are relevant to the items, but it produces comparatively good fit (de la Torre and Douglas, 2004, 2005). An object-oriented programming language, Ox program code (Doornik , 2002) using an EM algorithm is available for the DINA analysis.

CHAPTER 3

METHOD

Participants

The data were obtained from a randomly selected group of 2,993 eighth grade students enrolled in public schools in a Midwestern state. The entire sample was used for the applications of each model.

Testing Materials

The test consists of 86 eighth-grade math items, each of which has four response options. According to the test blueprint, the math test represents a hierarchical structure in which the superior level involves four standards, and the subordinate levels consist of benchmarks within the standards and indicators within the benchmarks. Figure 1 illustrates the hierarchical structure of standards, benchmarks, and indicators. The four standards represent Number and Computation, Algebra, Geometry, and Data; each standard includes two or more specific benchmarks. This study included only the four standards as dimensions for analysis. The test consists of 28 Number and Computation items, 23 Algebra, 18 Geometry, and 17 Data items as specified in the blueprint.

Although the items were originally developed for one of four standards according to the test blueprint, the items were later rescored for multiple indicators by a panel comprised of a mathematician, math educators, and an educational psychologist. If an item involves indicators from a standard other than the blueprint standard, it was rescored as including multiple standards. These scores were also represented in the augmented **Q-** matrix which shows the relationship of standards and specific items. According to the **Q-**

matrix based on the augmented blueprint, of the 86 items, 41, 46, 25, and 24 items were involved in the number, algebra, geometry, and data dimensions, respectively. For analysis, all the item responses were recoded into two categories, 0 for wrong answers and 1 for right answers.

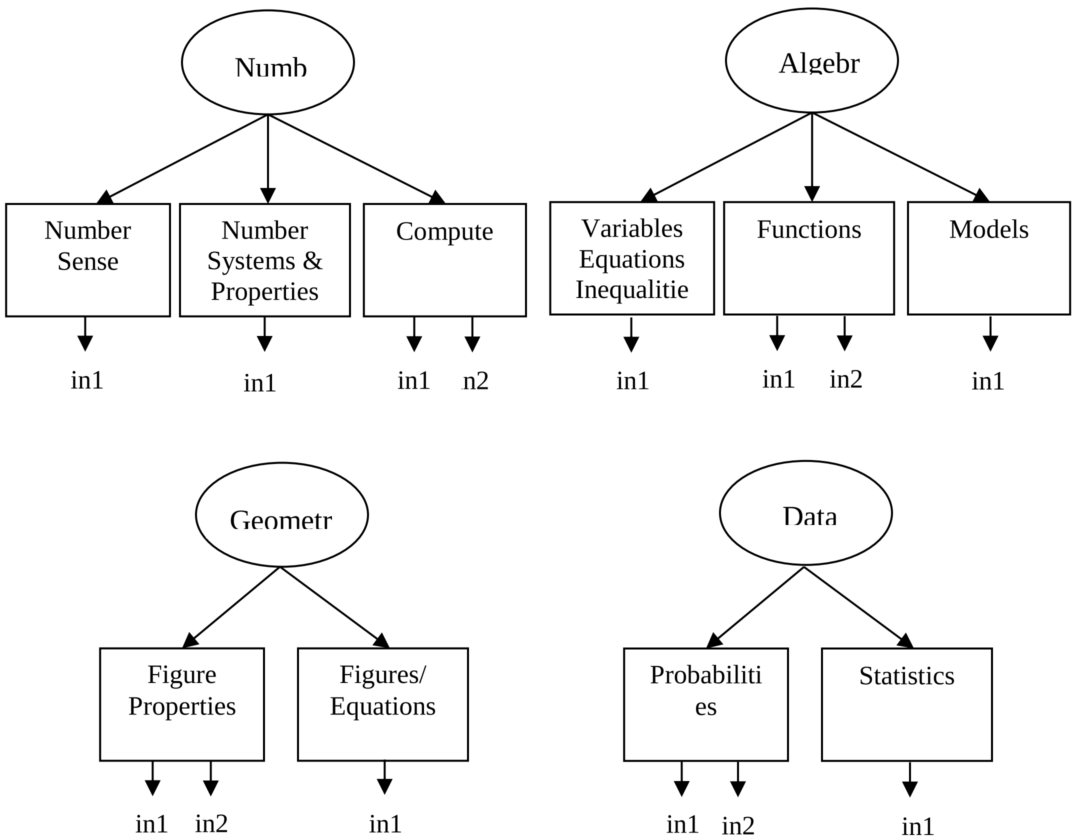

Figure1. *Hierarchical Structure of Standards, Benchmarks, and Indicators on a Math Test*.

Procedure

The 86 test items were divided into three sections, and only one section was administered online per day. Therefore, it took three days for students to complete all three parts; Parts 1, 2, and 3 included 30, 27, and 29 items, respectively.

Testing models

The EMIRT, full information bi-factor, traditional unidimensional IRT, and CMIRT models were tested using IRTPRO 2.1 (Thissen & Cai, 2011). The confirmatory MIRT models were based on both the test blueprint and the augmented test blueprint with the entire sample. Then, the MLTM-D was examined using SAS PROC NLMIXED for item parameters and an SPSS macro for person parameters. The interpretability of the dimensions, the reliability of the score estimates, and the fit between the models were examined and compared to identify which model fits the best. The current study also includes delineating a) dimension probabilities, and b) dimension pattern probabilities using alternative cutlines for the MIRT model. Lastly, the resulting data are compared with attribute pattern probabilities of the DINA model and then with component probabilities of the MLTM-D. In this paper, the term of dimension in the MIRT model will be interchangeable with the attribute in the DINA model, and the component in the MLTM-D.

CHAPTER 4

RESULTS

Descriptive Statistics

First, the descriptive statistics for item responses were obtained from the entire sample. The mean and the standard deviation of item responses were 60.7 and 15.1, respectively, and the average proportion of passing total items was 0.71, showing that the items were moderately easy. The skewness and the kurtosis of item responses were -0.409 and -0.726, respectively. The results for the item analysis based on classical test theory (CTT) are shown in Table A1. The column of the percent correct and the point-biserial correlation in the table can be interpreted as indices of item easiness and item discrimination, respectively. For example, if most students answered an item correctly, the percent correct would be high, indicating that the item is easy. Item 1 in part 1 was answered correctly by 2,867 (96%) out of 2,993 students, and the results can be explained by low item difficulty. On the other hand, if most students missed an item, the percent correct would be low, indicating that the item is difficult. Item number 24 in part 2 was correctly answered by only 987 (33%) out of 2,993 students, which indicates high item difficulty. In turn, the point-biserial correlation indices represent the relationship between an item score and total test score that examinees owe, representing how well an item can discriminate between students with high and low abilities. Items with a high point-biserial correlation are considered to discriminate between students with low and high abilities better than items with a low point-biserial correlation. Most items in the current test seemed to be fairly well discriminating, excluding item 4 and 86, with point-biserials exceeding 0.30.

The EMIRT Model

In order to explain the inter-correlations among the items and to roughly examine the data structure, a four-factor 2PL EMIRT model was fitted to the data. The parameters in the EMIRT model were estimated using IRTPRO 2.1. For estimation of item factor loadings, the Bock-Aitkin expectation maximization (BAEM) method was used, and for factor scores, expected a priori (EAP) estimation was used, with nine quadrature points. Table A2 shows the item estimates from the exploratory MIRT model. The rotated results of factor loadings were obtained by using the orthogonal CF-Varimax rotation method that IRTPRO 2.1 offers, and are provided in Table A3. The results of the rotated factor loadings showed that the data generally supported four dimensions. However, most items were loaded strongly on one dimension, which appeared to be a general math ability dimension, and the other three dimensions could not be clearly related to item content as represented in the test blueprint. Because the general dimension was dominant, the data structure may be unidimensional rather than multidimensional. Thus, unidimensional IRT and full information bifactor models should have been applied to the data and statistically compared to determine whether the data consist of one or multiple dimensions.

Unidimensional IRT Model

One-, two-, and three-parameter logistic models (i.e., 1PL, 2PL, 3PL models) and a three-parameter logistic model with common asymptote (3PLC) were analyzed. Appendix tables A4 through A7 include item estimates from the 1PL, 2PL, 3PL, and 3PLC models, respectively. The item parameters from the unidimensional IRT models were estimated using the BAEM method, and person scores were estimated using expected *a priori* (EAP) estimation, with the default number of 45 quadrature points. IRTPRO 2.1 provided three types of overall goodness-of-fit statistics: the -2loglikelihood, Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The -2loglikelihood values can be used for model comparison because the -2loglikelihood difference is asymptotically distributed chi-square with the degrees of freedom of difference. The results of the comparisons among the unidimensional IRT models are presented in Table 1, showing that the 3PL model best fit the data. IRTPRO also offers the indices of item-level misfit (S-X^2). The item-level fit indices from different four IRT models are provided in Tables A8 through A11, showing that most items were fitted well to all the unidimensional IRT models, except for the 1PL model.

Table 1. *Goodness of Fit Comparison among IRT Models*

	Model(# of parm)	**AIC**	**-2Log-Likelihood**	**Δ -2 Log-Likelihood**
Unidimensional IRTmodel	1PL(87)	242774.68	242600.68	
	2PL(172)	239178.45	238834.45	**3766.23***
	3PL Common(173)	238336.95	237990.95	**843.50***
	3PL(258)	238081.37	237565.37	**425.58***
Multidimensional IRT model	Bifactor model(258)	236,538.20		
	Original Blueprint-based CMIRT(172)	248,867.52	249,211.52	
	Augmented Blueprint-based CMIRT(222)	246,566.77	247,010.77	**2,300.75***

Full Information Bifactor Model

Unlike the traditional bifactor model, which is based on the correlation coefficients, the full information bifactor model works directly on the item response pattern. For this study, the full information bifactor model containing one general dimension and four specific dimensions was analyzed. A diagram representing this model is displayed in Figure 2. For the full information bifactor model, all items load on a general factor representing general mathematics ability, and also load on a specific factor to which each item is related. For item parameter estimation, marginal maximum likelihood estimation (MMLE) method based on the expectation-maximization (EM) algorithm was used, and person estimates were obtained via EAP estimation, with nine quadrature points. Table A12 shows the factor loadings, indicating that most items had high loadings on the first dimension and only a few items had high loadings on the secondary dimensions, which were defined by the four standards. Even though some items seem to have negative loadings on the secondary dimensions, indicating negative conditional dependencies, given the first dimension, still many items had positive loadings, representing that the test items had positive conditional dependencies on the specific dimensions. (e.g. See Appendix 13 for detailed information on item parameter estimates)

In a comparison with the 2PL unidimensional IRT model, the bifactor model showed a better fit with a chi-square difference of 2,296 (d.f. = 86, $p < 0.01$), representing that the math test would probably consist of more than one dimension. Specifically, the -2loglikelihood value from the unidimensional IRT analyses was 238,834 with 172 parameters, and that from the bifactor model was 236,538 with 258 parameters. Meanwhile, the empirical reliability estimate on the general factor was 0.93, which was high, showing that θ_g on the general factor are reliable. The specific factors had low reliability for the bifactor model, with values ranging from 0.38 to 0.49. As a result of the low reliability of specific dimensions, this model has limited diagnostic use.

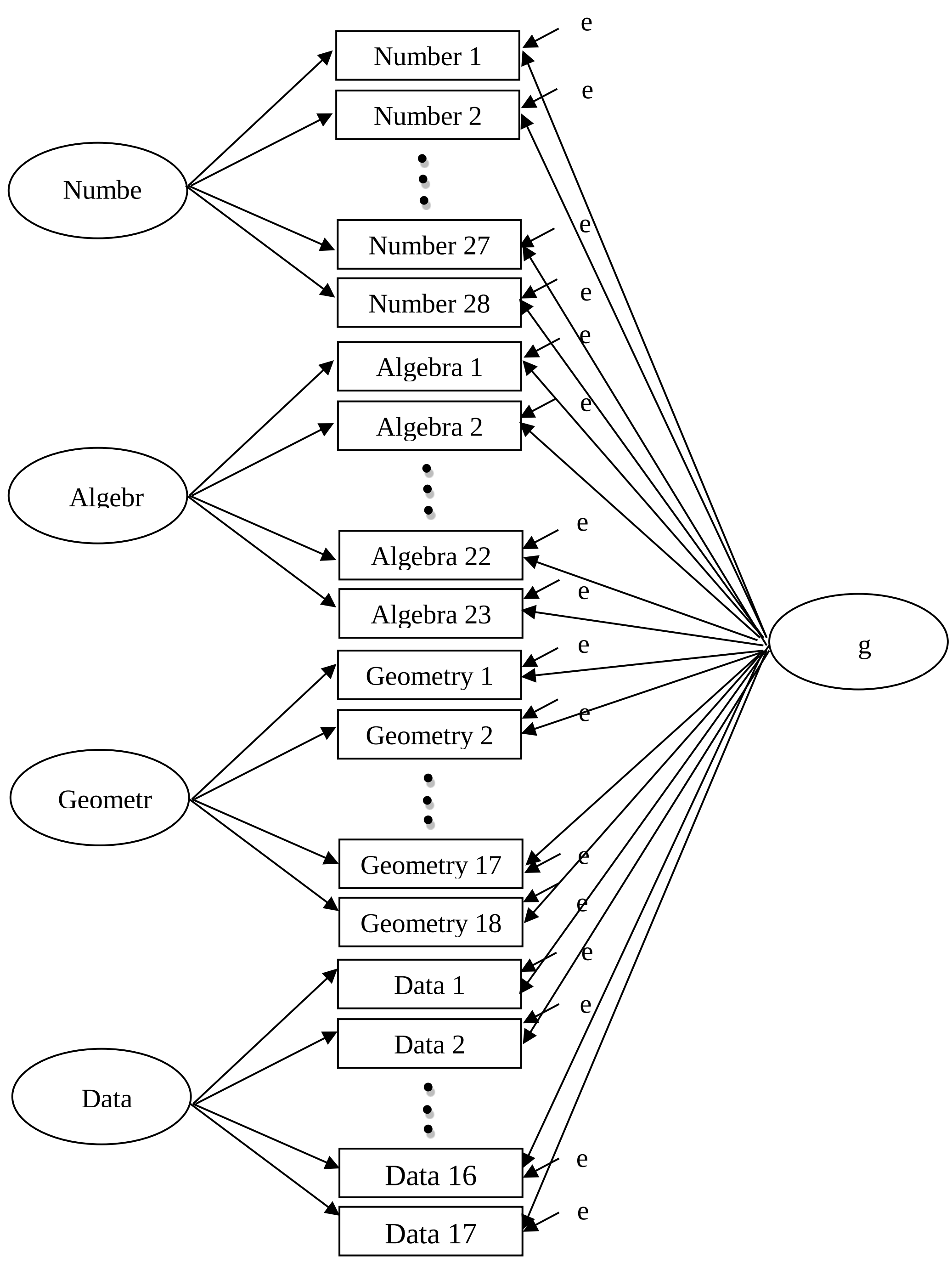

Figure 2. *Configuration of Bifactor Model*

The CMIRT Model

The application of the four factor CMIRT models were based on two different hypotheses of data structure: the original and augmented test blueprints. The corresponding diagrams of the models are provided in Figures 3 and 4. In Figure 3, items load on only one relevant dimension as indicated in the original blueprint; in Figure 4, items are split-loaded on more than one dimension. For item parameter estimations in both models, the BAEM method was used, and for person parameter estimation, the EAP method was used, with nine quadrature points. Tables A14 and A15 show the component weights of the CMIRT models based on both the original and the augmented test blueprint, respectively. In the model based on the original test blueprint, most of the dimensions seem to be influential, exceeding 0.6, with the exception of a few items. For example, item 4 and item 86 showed very small weights of -0.28 and 0.37 in relevant dimensions, respectively, indicating that the corresponding dimensions are less important or negatively influential to solving the item. Meanwhile, in the CMIRT model based on the augmented test blueprint, the weights of the dimensions also appeared to be significantly influential, but the effects differed from those of the previous model. Some items had relatively lower weights in specific dimensions when several dimensions defined the items. Both models fit the data quite well, as shown in Table 1, but the -2loglikelihood difference showed that the augmented test blueprint-based CMIRT model had a better fit with the Δ-2loglikelihood of 2,300.75 and the $\Delta d.f.$ of 50, $p < 0.01$. That is, the augmented blueprint-based CMIRT model explained the mathematics data better than the test blueprint-based CMIRT model.

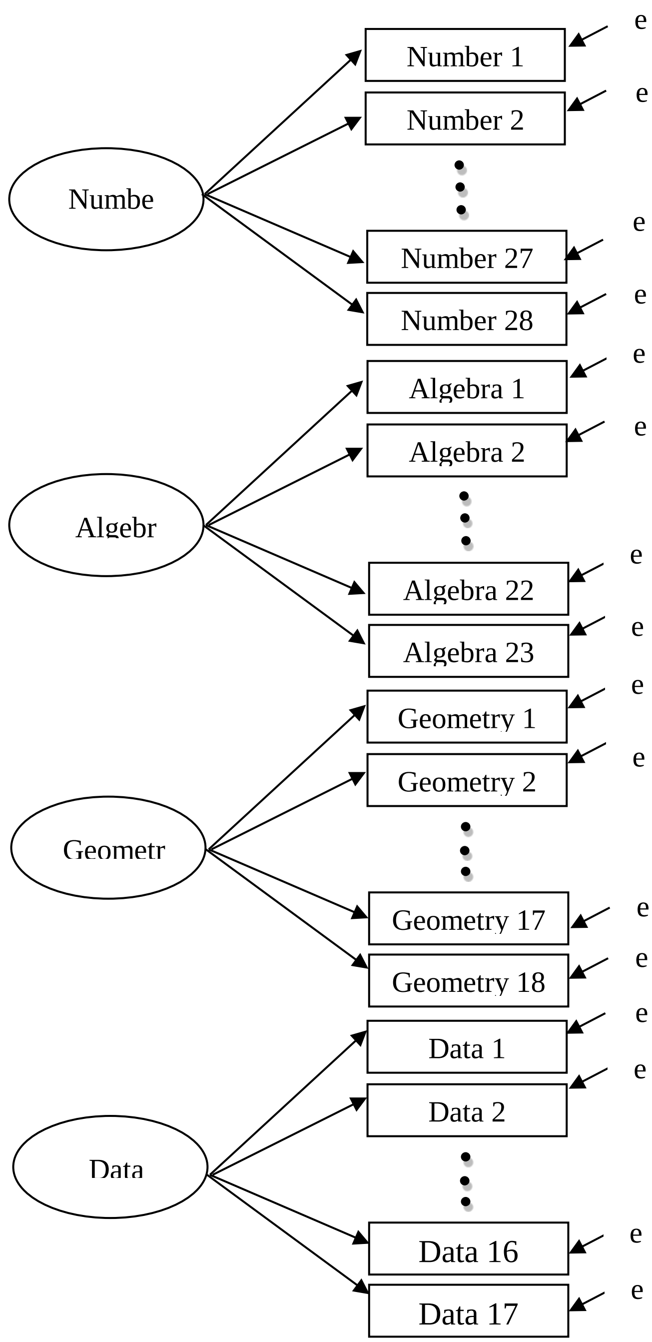

Figure 3. *Configuration of CMIRT Model Based on the Original Test Blueprint*

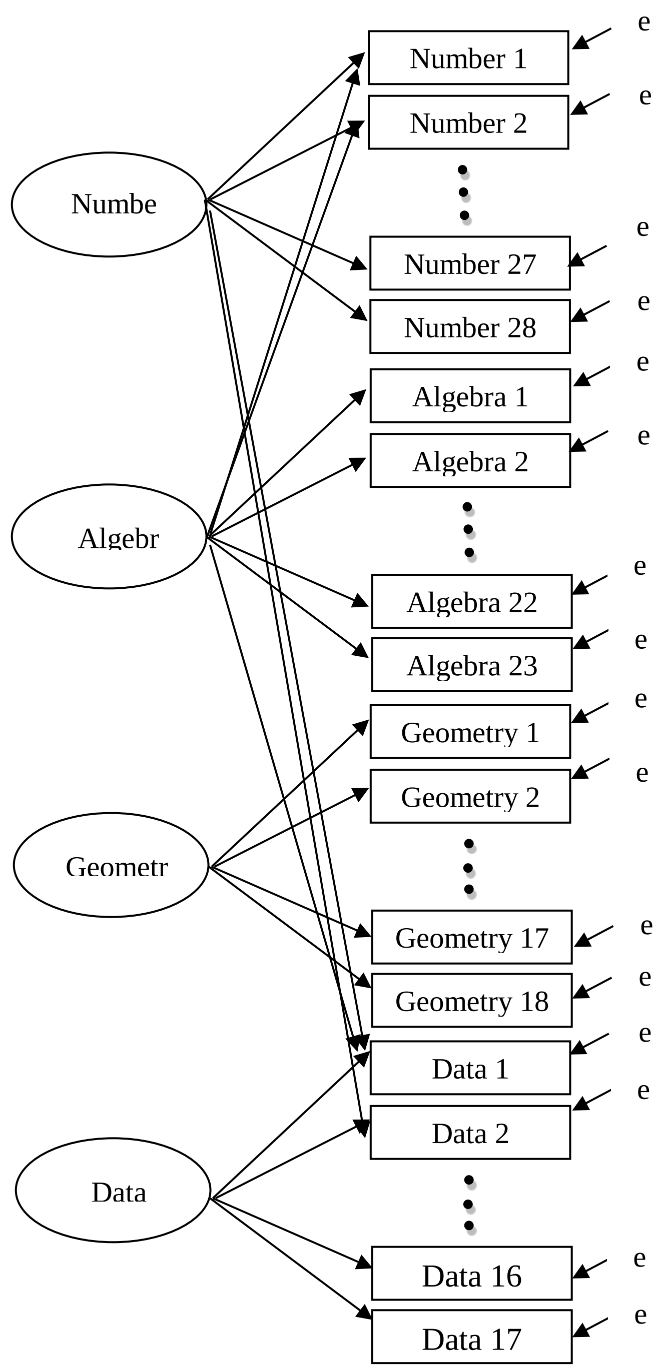

Figure 4. *Configuration of CMIRT Model Based on the Augmented Test Blueprint*

Item fit was assessed by S-χ^2, presented in Tables A16 and A17, of the 86 items a total of 34 and 36 misfit under the original and augmented blueprint based CMIRT models, respectively. Although the results of item-fit seemed to show that many items do not fit well to the model, this concern is negligible if one considers that the S-χ^2 test is very sensitive to the sample size (Orlando & Thissen, 2000; Stone & Zhang, 2003), which in this case is large.

The CMIRT analysis produces θ estimate for each dimension. Based on the fit comparison, descriptive statistics of the theta estimates for the augmented blueprint-based CMIRT model alone were observed (Table 2). The means and the standard deviations of theta estimates were close to 0 and 1, respectively, and the distribution graphs for the four dimensions indicate that they are likely normally distributed. The graphs in Figure 5 plot the theta distribution for the four dimensions. The empirical reliabilities of the thetas were computed based on the competency level estimates and error variances, the results of which are also represented in Table 2.

Table 2. *Descriptive Statistics of Theta Estimates for Four Components from Augmented Blueprint based CMIRT model*

	Theta1	Theta2	Theta3	Theta4
Min	-1.747	-1.837	-1.517	-1.524
Max	1.898	2.098	2.431	1.719
Mean	0.113	0.138	0.241	0.158
SD	0.698	0.770	0.717	0.621
MSE	0.327	0.298	0.219	0.196
Empirical Reliability	0.598	0.666	0.701	0.663

The empirical reliabilities ranged from 0.598 to 0.701, indicating that the reliabilities of the estimates of examinees' competency in each dimension are moderately high.

Specifically, the theta estimates were relatively less reliable on the Number and Computation dimension, and more reliable on the Geometry dimension.

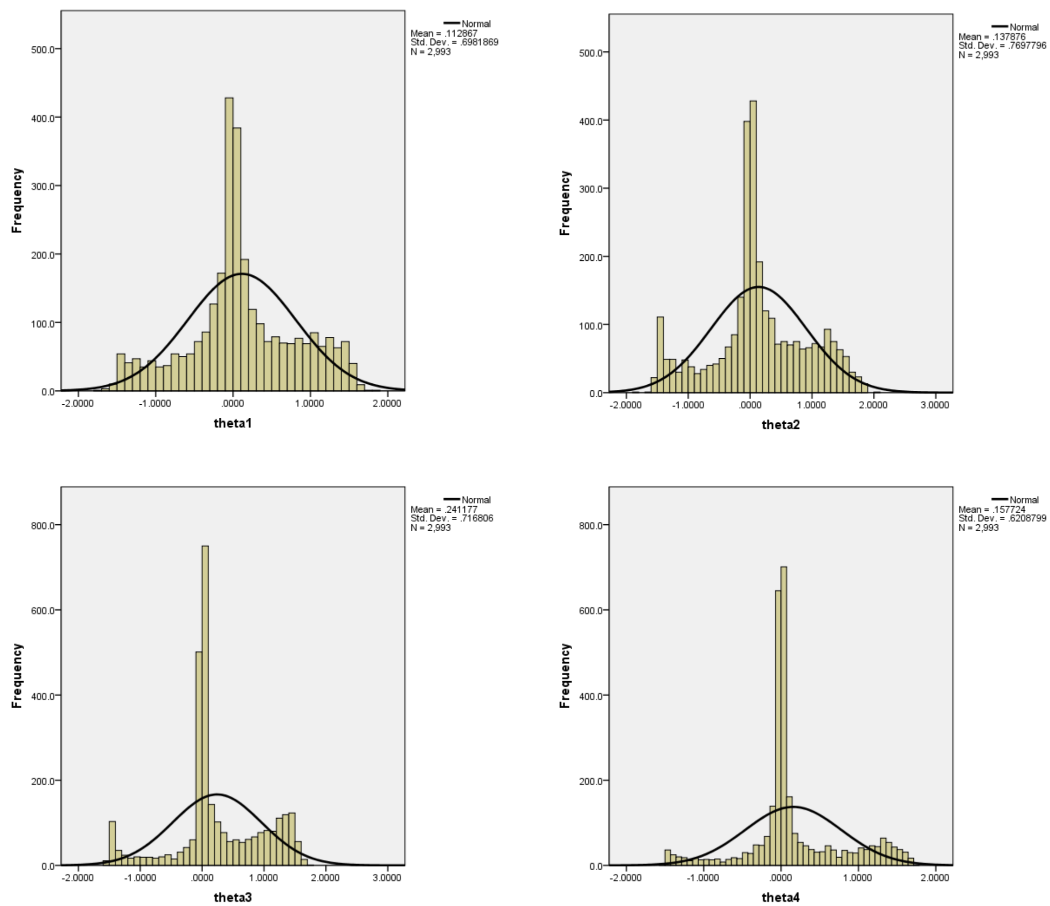

Figure 5. *Distribution Graphs of the Theta Estimates from the Augmented Blueprint Based CMIRT Model*

Correlations among the estimates of the dimensions were moderately low, ranging from 0.305 to 0.411, which is evidence that the four dimensions are not redundant. The correlations of the theta estimates are shown in Table 3.

Table 3. *Correlation Coefficients among Theta Estimates from Augmented Blueprint based CMIRT Model*

	Theta1	Theta2	Theta3	Theta4
Theta1 (Number)	1.00	0.36	0.41	0.37
Theta2 (Algebra)		1.00	0.41	0.41
Theta3 (Geometry)			1.00	0.31
Theta4 (Data)				1.00

The CMIRT models can estimate examinees' proficiency scores on each dimension, permitting diagnosis at the dimension level. Observing students' ability patterns of dimensions at or below a specific overall proficiency cutline is very important for instructional intervention. For example, a student with low proficiency on a specific dimension can be instructed so as to increase knowledge in that dimension, resulting in an improvement of the student's general performance. Four students with the same overall ability estimate were selected, and their dimension competency patterns were observed. The four students exhibited very different patterns of competency on each dimension (Figure 6). Student 2564 showed high competency on the Number and Computation dimension but low competency on the Algebra and the Geometry dimensions; student 969 showed high competency on the Geometry dimension, but showed low competency on the Algebra and the Data dimensions; student 1911 showed moderately high competency on all dimensions except for the Number and Computation dimension; finally, student 321 was moderately competent across all four standards without being very low or very high on any one. Given a specific competency level based on general performance; these four students differ in the dimensions on which their performance level is above or below the proficiency level, resulting in different remedial interventions. For example, the student 2564 would focus on the Algebra and Geometry dimensions, and the others would receive instruction in their own weaker dimensions.

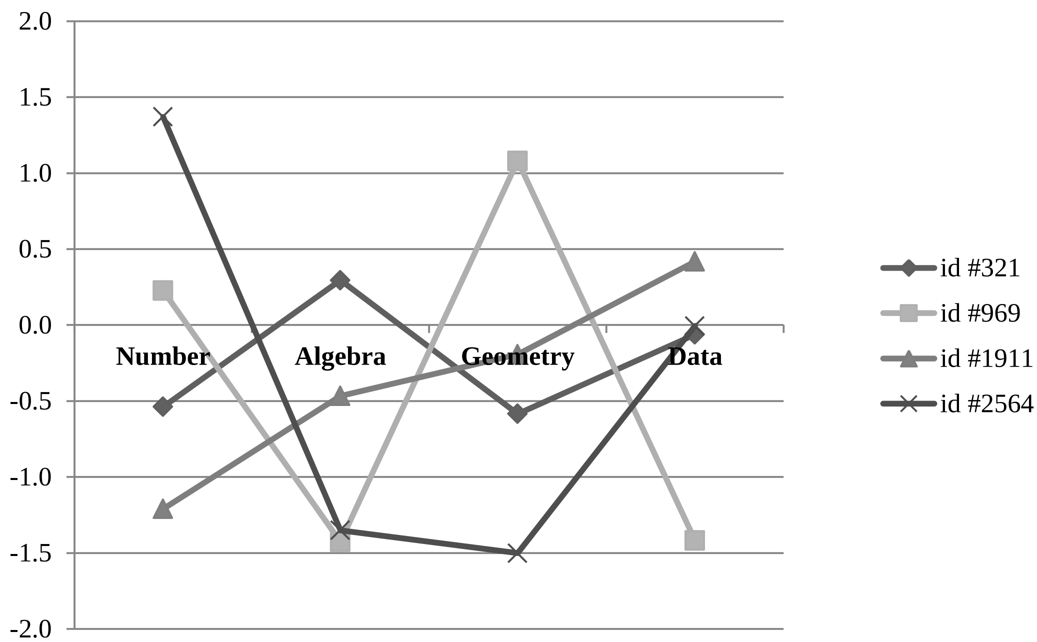

Figure 6. *Patterns of CMIRT Theta Estimates on Four Components from Students with Same Total Score*

Diagnosis from the CMIRT Model

Dimension and Pattern Probabilities

Because the CMIRT models offer item and person estimates for each dimension, it enables the computation of dimension probabilities. The dimension probabilities represent probabilities that each student will correctly answer items with a specific components based on the model parameters. The item-solving probabilities for each dimension can be calculated using the 2PL IRT formula with the corresponding trait score (θ_{sm}), dimension weight ($\alpha_{i(m)}$), and item easiness (δ_i). The formulas used for computing dimension probabilities are below:

$$P(X_{si_{(m)}} \mid \theta_{sm}, \delta_i, \alpha_{i_{(m)}}) = \frac{\exp(\alpha_{i_{(m)}} \theta_{sm} + \delta_i)}{1 + \exp(\alpha_{i_{(m)}} \theta_{sm} + \delta_i)} \text{ , and}$$

$$P(X_{sm}) = \frac{\Sigma P(X_{si_{(m)}})}{\Sigma i_{(m)}}.$$

For example, given the dimension weights ($\alpha_{i(m)}$) for the 41 items in the Number and Computation dimension and relevant person estimates (θ_{sm}), the item probability on the dimension that a student would answer each item correctly is computed. In turn, dimension probabilities that the student would master the Number and Computation dimension are obtained by averaging all item probabilities over the 41 items relevant to the dimension.

For the computation, SPSS 19.0, and Excel were used. Descriptive statistics of the dimension probabilities are represented in Table 4. The mean dimension probabilities ranged from 0.753 to 0.808, showing generally high competency across four dimensions, and the standard deviations ranged from 0.059 to 0.088, showing that dimension probabilities are not scattered largely across students. The measures of skewness showed that the algebra and the geometry dimensions were negatively skewed, but the other dimensions appear to be normally-distributed. The average dimension probabilities, p = 0.76, are higher than the proportion of passing items, p = 0.71, based on classical test theory.

Table 4. *Descriptive Statistics of Dimension Probabilities from Augmented Blueprint based CMIRT model*

Dimension (# of Items)	N	Mean	SD	Skewness	Kurtosis
Number (41)	2993	0.756	0.068	-0.521	0.187
Algebra (46)	2993	0.755	0.082	-0.796	0.368
Geometry (25)	2993	0.808	0.088	-1.505	2.608
Data (24)	2993	0.753	0.059	-0.305	1.329

This difference in probabilities is likely due to the estimation method used for item parameters. The BAEM method assumes the normal distribution for person scores for estimating marginal maximum likelihood. The skewness of two dimensions, along with rescaling, can influence the higher probabilities computed from the model parameters than from the classical test theory overall p-value.

Dimension probabilities also provide diagnostic information for instructional intervention in the same manner as theta estimates. Figure 7 illustrates the pattern of dimension probabilities for four students: student 1944 has probably mastered Algebra and Geometry, but he or she probably needs supplementary learning on the Number and Computation topics; student 1335 has mastered on all dimensions excluding geometry; student 1258 has mastered algebra and data, has not mastered Number and Computation or Geometry; finally, student 969 showed non-mastery on the Algebra and Data dimensions, and showed mastery of the Number and Computation and Geometry dimensions.

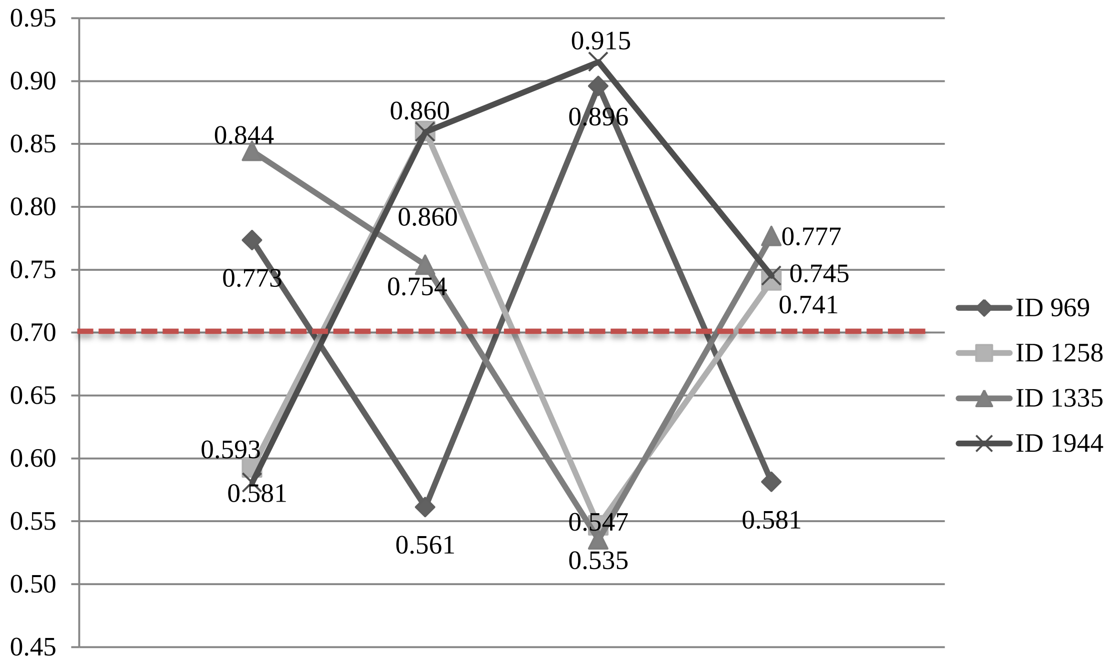

Figure 7. *Patterns of CMIRT Dimension Probabilities of Four Student Samples*

Although dimension probabilities and dimension estimates can provide diagnostic information, it may not be easy for students or instructors to interpret the profiles effectively. Setting a proficiency cutline for dimension probabilities can generate discrete patterns for each dimension, just as classification models do, and can reduce the complexity of interpretation. Table 5 shows pattern probabilities that were produced using different proficiency cutlines. A total of 2^4 = 16 possible patterns were generated from the four dimensions, and the distributions included in each pattern were different according depending on which cutline was used. In setting low cutlines between 0.58 and 0.63, approximately 80% of students showed mastery on three or four dimensions, and in setting a high cutline of 0.78, only 11.9% of students showed non-mastery on all dimension.

Table 5. *Distribution in Dimension Patterns at Different Dimension Proficiency Cutlines: Augmented Blueprint-based CMIRT and DINA.*

	Pattern	**Augmented Blueprint-based CMIRT**							**DINA**
		>= 0.58	**>= 0.63**	**>=0.68**	**>=0.70**	**>= 0.73**	**>= 0.74**	**>= 0.78**	
1	**1111**	2,618(87.47%)	2,257(75.41%)	1,963(65.59%)	1,840(61.48%)	1,543(51.55%)	1,211(40.46%)	355(11.86%)	1340 (44.77%)
2	**1110**	32(1.07%)	85(2.84%)	117(3.91%)	139(4.64%)	204(6.82%)	363(12.13%)	233(7.78%)	36 (1.20%)
3	**1101**	144(4.81%)	129(4.31%)	116(3.88%)	110(3.68%)	68(2.27%)	49(1.64%)	3(0.10%)	38 (1.27%)
4	**1011**	156(5.21%)	222(7.42%)	270(9.02%)	268(8.95%)	240(8.02%)	211(7.05%)	83(2.77%)	48 (1.60%)
5	**0111**	11(0.37%)	156(5.21%)	260(8.69%)	280(9.36%)	312(10.42%)	279(9.32%)	95(3.17%)	41 (1.37%)
6	**1100**	3(0.10%)	9(0.30%)	16(0.53%)	16(0.53%)	13(0.43%)	22(0.74%)	10(0.33%)	33 (1.10%)
7	**1010**	5(0.17%)	37(1.24%)	50(1.67%)	61(2.04%)	85(2.84%)	138(4.61%)	236(7.89%)	18 (0.60%)
8	**1001**	21(0.70%)	47(1.57%)	60(2.00%)	60(2.00%)	52(1.74%)	40(1.34%)	4(0.13%)	11 (0.37%)
9	**0110**	0(0.00%)	3(0.10%)	19(0.63%)	26(0.87%)	71(2.37%)	164(5.48%)	299(9.99%)	82 (2.74%)
10	**0101**	0(0.00%)	21(0.70%)	40(1.34%)	48(1.60%)	53(1.77%)	38(1.27%)	9(0.30%)	43 (1.44%)
11	**0011**	3(0.10%)	16(0.53%)	44(1.47%)	80(2.67%)	156(5.21%)	153(5.11%)	118(3.94%)	29 (0.97%)
12	**1000**	0(0.00%)	3(0.10%)	10(0.33%)	12(0.40%)	33(1.10%)	45(1.50%)	26(0.87%)	20 (0.67%)
13	**0100**	0(0.00%)	4(0.13%)	6(0.20%)	7(0.23%)	13(0.43%)	24(0.80%)	33(1.10%)	93 (3.11%)
14	**0010**	0(0.00%)	0(0.00%)	7(0.23%)	17(0.57%)	65(2.17%)	146(4.88%)	1133(37.85%)	42 (1.40%)
15	**0001**	0(0.00%)	4(0.13%)	14(0.47%)	22(0.74%)	53(1.77%)	53(1.77%)	25(0.84%)	23 (0.77%)
16	**0000**	0(0.00%)	0(0.00%)	1(0.03%)	7(0.23%)	32(1.07%)	57(1.90%)	331(11.06%)	1096 (36.62%)
	Total	2993(100%)	2993(100%)	2993(100%)	2993(100%)	2993(100%)	2993(100%)	2993(100%)	2993 (100%)
Number of Exa minee (%)		Number of Examinees Involved in the Same Pattern Groups From both DINA and Augmented Blueprint-based CMIRT Models							
		1343(44.87%)	**1352(45.17%)**	**1349(45.07%)**	**1343(44.87%)**	**1301(43.47%)**	**1181(39.46%)**	**776(25.93%)**	

However, in setting the cutlines between 0.68 and 0.73, approximately 50% of students showed mastery on all the four dimensions, and the distribution of the remaining 50% of students were widely spread out across all over the remaining possible patterns. Specifically, given a cutline of 0.73, though approximately 52% of students showed mastery on the four dimensions, the remaining students were widely involved in the patterns involving fewer than four dimensions. Of the students proficient on the same number of dimensions, they differ in terms of which specific dimensions they were proficient. The different patterns each require different instructional guidance. Defining cutlines too high or too low does not appear to be desirable for either diagnostic or instructional purposes. The following section contains comparison in dimension pattern probabilities with one of the classification models, the DINA model and the MLTM-D.

The DINA Model

The estimation of item parameters (i.e., guess and slip parameters) and attribute classification for the DINA model was done using specialized code written in the Ox program. The estimation procedures are based on the EM algorithm. The results include test- and item-level fit statistics as well as item estimates and attribute classification. The guess and slip estimates with the corresponding standard errors are shown in Table A8; the guess estimates ranged from 0.106 to 0.932, indicating that the math test consists of items with various levels of guessing. The mean and the standard deviation of the guess estimates were 0.541 and 0.212, respectively. It could be interpreted that students have the probability of a correct response of 0.54 on average, even if they are not proficient on all the required attributes to solve an item. The guess parameters of 20 items were empirically plausible, producing lower value than .35. The range of slip parameter estimates was between 0.002 and 0.658, and the mean and the standard deviation were 0.145 and 0.138, respectively. The mean of the slip parameter estimates does not appear highly large, representing that not many proficient students will incorrectly answer an item. Table 5 also includes the attribute patterns from the DINA model and the number of students involved in each pattern: most students appear to be classified into the two extreme patterns. That is, 44.8% of students (1,340 out of 2,993) were classified as possessing all the required attributes for solving the 8th grade math items, and 36.6% of students (1,096 out of 2,993) were classified as not possessing any attributes required for math problem solving. Only 18.6% of students were involved in the remaining patterns in which one or more attributes are missed.

Comparison of Dimension Patterns from both the DINA and CMIRT models

The comparison of the dimension patterns between the DINA and CMIRT model were examined by counting the number of students included in the same dimension patterns. The last row of Table 5 also shows the number of students who are in the matching groups in the dimension patterns. As previously noted, dimension probabilities of the CMIRT model provide different dimension patterns using the different cutlines. When setting proficiency cutlines between 58% and 70% for dimension probabilities of the CMIRT model, approximately 45% of students fell in the same dimension pattern groups. Tables 6 through 11 show the distribution of students involved in the corresponding dimension patterns based on both the CMIRT and DINA models. A low cutline of 58% included most students in the dimension pattern showing mastery on all four dimensions (i.e., 1111). All students who were classified to have mastered all four dimensions according to the DINA model were also classified as masters based on the

CMIRT model. However, because setting too low cutline positioned too many students (2,618 out of 2,993) at the perfect mastery pattern, the CMIRT pattern at 58% cutline were not appropriate to evaluate students at the other dimension classification from the DINA model. Specifically, the CMIRT patterns using 58% cutline seemed to be low for the purpose of differently classifying students in their mastery patterns. While setting a proficiency cutline of 70% or 73% for dimension probabilities scattered students in diverse patterns, still most students were classified into pattern groups of mastery in three or all dimensions. As with the proficiency cutline between 58% and 73%, approximately 45% of students were included in the same patterns by using both DINA and CMIRT models. That is, regardless of which proficiency cutline between 58% and 73% was used, the number of students included in the same attribute pattern was consistent. However, if considering up to the number of students showing difference in only an dimension pattern, the number of matching students could be increased up to 1,834 (61.3%). Setting a high proficiency cutline of 78%, relatively fewer number of students, 776 (25.93%), were included in matching groups, but the number of students with similar patterns was dramatically increased up to 2,087 (69.73%) if also considering the inconsistent patterns only in a dimension. Compared to the other cutlines, the cutline of 78% spread students out wider, and could predict similarly for students to show nonmastery in two or more attributes by the DINA model. Approximately 45% of students were classified into having same mastery patterns in all four attributes, and approximately 70% of students were classified to commonly have mastery patterns in three or four attributes. In summation, the CMIRT model generally seemed to provide dimension patterns comparable to the DINA model, but the degree of prediction was different according to the proficiency cutline used. The main difference between both models was that the DINA model placed students at extreme dimension patterns of possessing either all attributes or none of the attributes, whereas the CMIRT model classified them in more diverse patterns.

Table 6. *CMIRT Attribute Probabilities Using 58% Cutline*

		Attribute Patterns from the DINA Model																
		1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
Attribute Patterns from the Augmented Blueprint-based CMIRT Model	1111	**1,339**	36	37	45	41	30	17	11	80	41	25	20	88	38	19	751	2,618
	1110	1	**0**	0	0	0	0	0	0	2	0	0	0	0	0	0	29	32
	1101	0	0	**1**	0	0	3	0	0	0	2	0	0	5	0	0	133	144
	1011	0	0	0	**3**	0	0	1	0	0	0	4	0	0	4	3	141	156
	0111	0	0	0	0	**0**	0	0	0	0	0	0	0	0	0	1	10	11
	1100	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	0	3	3
	1010	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	5	5
	1001	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	21	21
	0110	0	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0
	0101	0	0	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0
	0011	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	0	3	3
	1000	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	0	0
	0100	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	0
	0010	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0
	0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0
	0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0
	Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,096	2,993

Total Number of Students in the Same Attribute Pattern Groups: **1,343(44.8%),**
Total Number of Students with difference in One or Less Attribute Pattern: **1,515(50.62%)**

Table 7. *CMIRT Attribute Probabilities Using 63% Cutline*

	Attribute Patterns from the DINA Model																
Attribute Patterns from the Augmented Blueprint-based CMIRT Model	1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
1111	**1,326**	34	35	36	31	29	16	9	60	33	23	19	76	32	13	485	2,257
1110	3	**2**	0	0	0	1	1	0	9	0	0	0	5	0	0	64	85
1101	0	0	**3**	0	0	3	0	0	0	5	0	0	5	0	3	110	129
1011	3	0	0	**12**	1	0	1	2	0	0	5	1	0	6	5	186	222
0111	8	0	0	0	**9**	0	0	0	13	5	1	0	7	3	2	108	156
1100	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	0	9	9
1010	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	37	37
1001	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	47	47
0110	0	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	3	3
0101	0	0	0	0	0	0	0	0	0	**0**	0	0	0	0	0	21	21
0011	0	0	0	0	0	0	0	0	0	0	**0**	0	0	1	0	15	16
1000	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	3	3
0100	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	4	4
0010	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0
0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	4	4
0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0
Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,096	2,993

Total Number of Students in the Same Attribute Pattern Groups: **1,352(45.2%),**
Total Number of Students with difference in One or Less Attribute Pattern: **1,560(52.12%)**

Table 8. *CMIRT Attribute Probabilities Using 68% Cutline*

		Attribute Patterns from the DINA Model																
		1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
Attribute Patterns from the Augmented Blueprint-based CMIRT Model	1111	**1,301**	28	33	30	24	28	12	9	45	28	18	16	66	17	10	298	1,963
	1110	13	**6**	0	0	0	2	1	0	13	0	0	0	6	4	0	72	117
	1101	0	0	**5**	0	0	3	0	0	0	5	0	1	7	0	3	92	116
	1011	8	0	0	**17**	1	0	5	2	1	0	8	3	0	10	5	210	270
	0111	17	2	0	1	**16**	0	0	0	23	7	3	0	13	9	4	165	260
	1100	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	0	16	16
	1010	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	50	50
	1001	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	60	60
	0110	1	0	0	0	0	0	0	0	**0**	0	0	0	1	0	0	17	19
	0101	0	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	37	40
	0011	0	0	0	0	0	0	0	0	0	0	**0**	0	0	2	1	41	44
	1000	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	10	10
	0100	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	6	6
	0010	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	7	7
	0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	14	14
	0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	1
	Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,096	2,993

Total Number of Students in the Same Attribute Pattern Groups: **1,349(45.1%)**,
Total Number of Students with difference in One or Less Attribute Pattern: **1,615(54.0%)**

Table 9. *CMIRT Attribute Probabilities Using 70% Cutline*

		Attribute Patterns from the DINA Model																
		1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
Attribute Patterns from the Augmented Blueprint-based CMIRT Model	1111	**1,269**	25	32	23	20	23	10	9	39	24	16	16	64	16	7	247	1,840
	1110	28	**8**	0	0	0	4	1	0	15	0	0	0	7	4	0	72	139
	1101	0	0	**5**	0	0	5	0	0	0	5	0	1	7	0	4	83	110
	1011	13	0	1	**23**	2	0	4	2	1	2	7	3	0	10	5	195	268
	0111	28	3	0	2	**19**	1	1	0	23	8	4	0	12	9	6	164	280
	1100	0	0	0	0	0	**0**	0	0	0	0	0	0	0	0	0	16	16
	1010	0	0	0	0	0	0	**2**	0	0	0	0	0	0	0	0	59	61
	1001	0	0	0	0	0	0	0	**0**	0	0	0	0	0	0	0	60	60
	0110	2	0	0	0	0	0	0	0	**4**	0	0	0	2	0	0	18	26
	0101	0	0	0	0	0	0	0	0	0	**4**	0	0	1	0	0	43	48
	0011	0	0	0	0	0	0	0	0	0	0	**2**	0	0	3	1	74	80
	1000	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	12	12
	0100	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	7	7
	0010	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	17	17
	0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	22	22
	0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**7**	7
	Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,09 6	2,993

Total Number of Students in the Same Attribute Pattern Groups: **1,343(44.9%),**
Total Number of Students with difference in One or Less Attribute Pattern: **1,655(55.3%)**

Table 10. *CMIRT Attribute Probabilities Using 73% Cutline*

		Attribute Patterns from the DINA Model																
		1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
Attribute Patterns from the Augmented Blueprint-based CMIRT Model	1111	**1,171**	15	27	16	14	17	4	5	22	19	6	13	42	10	5	157	1,543
	1110	70	**15**	1	0	0	10	3	0	17	0	0	0	9	2	0	77	204
	1101	0	0	**7**	0	0	3	0	0	0	5	0	1	6	0	2	44	68
	1011	25	0	2	**25**	3	0	5	6	2	1	11	6	0	6	4	144	240
	0111	64	3	1	2	**23**	1	1	0	25	11	8	0	29	9	6	129	312
	1100	0	0	0	0	0	**2**	0	0	0	0	0	0	0	0	0	11	13
	1010	2	2	0	1	0	0	**3**	0	1	0	1	0	0	5	0	70	85
	1001	0	0	0	0	0	0	0	**0**	0	1	0	0	0	0	1	50	52
	0110	7	1	0	0	0	0	0	0	**13**	0	0	0	5	2	0	43	71
	0101	0	0	0	0	0	0	0	0	0	**6**	0	0	2	0	1	44	53
	0011	0	0	0	4	1	0	0	0	2	0	**3**	0	0	7	4	135	156
	1000	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	0	33	33
	0100	0	0	0	0	0	0	0	0	0	0	0	0	**0**	0	0	13	13
	0010	1	0	0	0	0	0	2	0	0	0	0	0	0	**1**	0	61	65
	0001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**	53	53
	0000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**32**	32
	Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,096	2,993

Total Number of Students in the Same Attribute Pattern Groups: **1,301(43.5%),**
Total Number of Students with difference in One or Less Attribute Pattern: **1,834(61.3%)**

Table 11. *CMIRT Attribute Probabilities Using 78% Cutline*

Attribute Patterns from the Augmented Blueprint-based CMIRT Model		**Attribute Patterns from the DINA Model**																
		1111	1110	1101	1011	0111	1100	1010	1001	0110	0101	0011	1000	0100	0010	0001	0000	Sum
	1111	**353**	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	355
	1110	226	**3**	2	0	0	2	0	0	0	0	0	0	0	0	0	0	233
	1101	1	0	**1**	0	0	0	0	0	0	1	0	0	0	0	0	0	3
	1011	79	0	1	**2**	0	0	0	0	0	0	0	0	0	0	0	1	83
	0111	88	0	0	0	**4**	0	0	0	0	1	1	0	0	0	0	1	95
	1100	0	0	3	0	0	**4**	0	0	0	0	0	1	2	0	0	0	10
	1010	154	8	4	7	0	8	**3**	4	0	1	2	6	2	4	0	33	236
	1001	0	0	2	0	0	0	0	**0**	0	0	0	0	0	0	0	2	4
	0110	199	9	5	0	10	6	0	0	**23**	5	0	0	25	1	0	16	299
	0101	1	0	1	0	0	0	0	0	0	**5**	0	0	0	0	1	1	9
	0011	61	0	3	11	6	0	0	3	1	4	**3**	0	0	0	7	19	118
	1000	0	0	0	0	0	0	0	0	0	1	0	**2**	0	0	0	23	26
	0100	0	0	2	0	0	4	0	0	0	5	0	0	**7**	0	0	15	33
	0010	178	16	9	28	21	8	15	4	58	18	23	10	55	**37**	9	644	1133
	0001	0	0	1	0	0	0	0	0	0	2	0	0	0	0	**5**	17	25
	0000	0	0	2	0	0	1	0	0	0	0	0	1	2	0	1	**324**	331
	Sum	1,340	36	38	48	41	33	18	11	82	43	29	20	93	42	23	1,096	2,993

Total Number of Students in the Same Attribute Pattern Groups: **776(25.93%),**
Total Number of Students with difference in One or Less Attribute Pattern: **2,087(69.73%)**

The similarity of the dimension patterns between the DINA and CMIRT model were also observed by correlation coefficients, φ. Table 12 shows the φ coefficients between all the possible patterns from both models. The results revealed that the pairs of the patterns acquired from both the CMIRT and DINA models were moderately correlated in most cases.

Table 12. *Phi Coefficient Coefficients between Attribute Patterns from the DINA and CMIRT Models*

		DINA			
		Number	Algebra	Geometry	Data
CMIRT 58%	Number	0.07*	0.08*	0.08*	0.06*
	Algebra	0.25*	0.30*	0.25*	0.24*
	Geometry	0.24*	0.25*	0.27*	0.25*
	Data	0.11*	0.12*	0.11*	0.12*
CMIRT 63%	Number	0.26*	0.20*	0.20*	0.22*
	Algebra	0.32*	0.40*	0.32*	0.31*
	Geometry	0.27*	0.28*	0.31*	0.27*
	Data	0.21*	0.19*	0.20*	0.23*
CMIRT 68%	Number	0.36**	0.28**	0.28**	0.30**
	Algebra	0.73**	0.47**	0.37**	0.37**
	Geometry	0.30**	0.30**	0.34**	0.29**
	Data	0.24**	0.22**	0.22**	0.27**
CMIRT 70%	Number	0.39**	0.31**	0.30**	0.33**
	Algebra	0.39**	0.50**	0.39**	0.39**
	Geometry	0.31**	0.31**	0.35**	0.30**
	Data	0.23**	0.21**	0.21**	0.27**
CMIRT 73%	Number	0.27**	0.37**	0.36**	0.39**
	Algebra	0.45**	0.58**	0.44**	0.44**
	Geometry	0.33*	0.33**	0.38**	0.31**
	Data	0.26**	0.25**	0.24**	0.33**
CMIRT 78%	Number	0.56**	0.46**	0.47**	0.50**
	Algebra	0.53**	0.58**	0.50**	0.52**
	Geometry	0.38**	0.39**	0.45**	0.37**
	Data	0.40**	0.36**	0.37**	0.45**

However, when setting a cutline of 58% for dimension probabilities, the φ coefficients for the Number and Computation and Data dimensions were 0.07 and 0.12, which was very low, indicating that those two sets of dimension patterns were not related. The highest φ coefficients between both models were obtained when setting the proficiency cutline to 78%, in which case the φ correlations ranged from 0.45 to 0.58, which are moderately high.

The MLTM-D

The MLTM-D, a more recent diagnostic model, was analyzed. Item parameters were estimated using a SAS macro designed based on the marginal maximum likelihood estimation (MMLE) method. The item difficulty and person parameter on each component were initialized to zero and one, respectively. The mean item difficulty on each component was relatively low, implying that items are easy on average. The standard deviation on each component was moderately high, indicating that the item difficulty within each component varied over a wide range. The difficulty and descriptive statistics of the MLTM-D item parameters are presented in Table A19 and Table 13, respectively.

Table 13. *Descriptive Statistics for MLTM-D Component Item Estimates*

	Component 1 (Number)	Component 2 (Algebra)	Component 3 (Geometry)	Component 4 (Data)
Min	-6.08	-6.50	-5.76	-6.50
Max	1.70	1.35	1.29	1.92
Mean	-2.40	-1.57	-1.85	-2.36
SD	1.86	1.45	1.55	2.20

For estimating the person parameter on each component, the SPSS macro based on the expected a posteriori (EAP) estimation was used. The descriptive statistics of the component thetas are shown in Table 14: the mean theta for each component was closed to 0, and the standard deviations were a little larger than 1. The large standard deviations denote that the students' thetas on each component are dispersed. Additionally, the empirical reliability on each component ranges from 0.77 to 0.85, indicating reliable measurement of each component theta. Finally, the correlations among the four components range from 0.63 to 0.75, which are moderately high. The correlation coefficients are found in Table 15.

Table 14. *Descriptive Statistics of Theta Estimates for Four Components from MLTMD*

	Theta1	Theta2	Theta3	Theta4
Min	-2.37	-2.38	-2.44	-2.41
Max	2.15	2.21	1.88	2.01
Mean	0.06	0.08	0.09	0.08
SD	1.13	1.16	1.23	1.12
MSE	0.61	0.48	0.52	0.61
Empirical Reliability	0.77	0.85	0.85	0.77

Table 15. *Correlation Coefficients among Theta Estimates from MLTMD*

	Theta1	Theta2	Theta3	Theta4
Theta1 (Number)	1.00	0.66	0.70	0.63
Theta2 (Algebra)		1.00	0.75	0.67
Theta3 (Geometry)			1.00	0.65
Theta4 (Data)				1.00

In order to assess item fit, predicted frequencies of correct item responses were obtained and compared to observed frequencies. Predicted item probabilities for each student using the MLTM-D item and person estimates were obtained, and then the predicted item probabilities were categorized into fixed intervals for each item. For students within each category, observed frequencies of correct item responses were also obtained and these relative probabilities were used to compute item goodness of fit with the predicted probabilities. Item fit was examined by a chi-square goodness of fit test. The chi-square tests provide measures of goodness fit by comparing discrepancies between the observed frequencies and the expected frequencies under the fitting model. The predicted probability was computed for each person on each item and then the probabilities were categorized into the fixed number of groups with similar probabilities. The number of categories used here was fifteen. The expected and observed frequencies of examinees answering correctly in each category on each were obtained, and chi-square value was computed across categories. Table A20 includes item fit for each item. 61 items were fitted well at the p level of 0.01, but the remaining items were working appropriately. Especially, item 4 in part 1 and item 29 in part 3 had very high chi-square value, implying large discrepancy between observed and expected frequencies in each category. These items were the ones that were proven to be working poorly through previous analysis. Figure 8 presents the regression of observed proportions passing each item based on the fifteen category groups. The correlation coefficient was 0.949, which is high. The result indicates that overall model fit seems to be appropriate.

Diagnosis from the MLTM-D

As mentioned earlier, observing patterns on components for students whose general math proficiency is below a fixed cutline is very informative and valuable for educational interventions. With a cutline of 58% (50 items correct out of 86 items total), the component competency levels of students below the cutline were examined. Although the students had same overall proficiency on math items, they are different in their respective component competencies. For example, the standard deviations among the four component competencies ranged from 0.35 to 0.62, indicating that component thetas between students with same general competency are moderately deviated. Figure 9 presents the example of the component competency patterns from four students with the same overall ability. Similar to the CMIRT model, the MLTM-D model provides different competency patterns based on which level of cutline was set to for the component thetas. Suppose that the cutline of component competency score was set at -0.70.

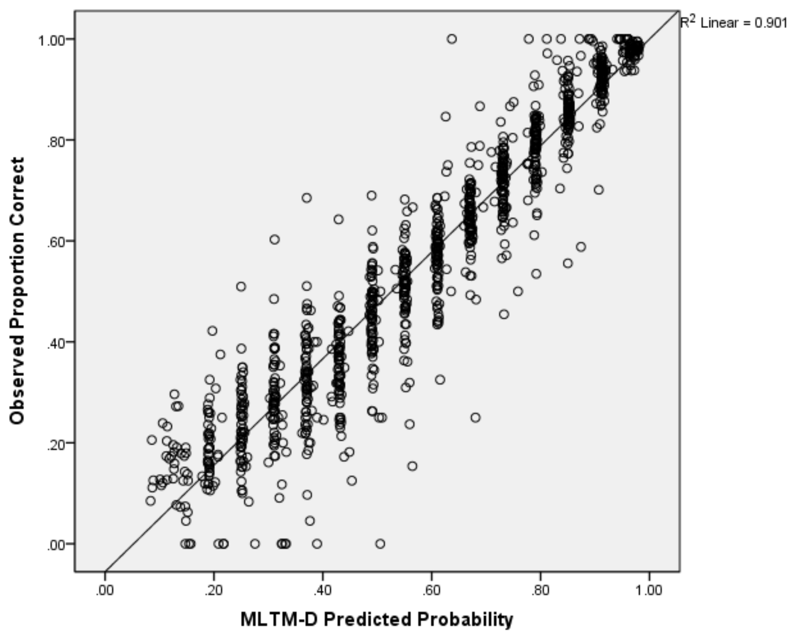

Figure 8. *Regression of Observed Proportion Passing Each Item on MLTM-D Prediction across Categories*

According to this cutline, the component patterns of the four students could be classified into two different groups such as competency vs. non-competency groups. Student 134 seemed to have mastery of the Algebra, Geometry, and Data components, but showed non-mastery in the Number and Computation component; student 243 showed mastery in the Number and Computation and Geometry components but did not have mastery of the

Algebra and Data components; finally, students 580 and 1646 showed mastery on all components except for the Algebra and Geometry components, respectively.

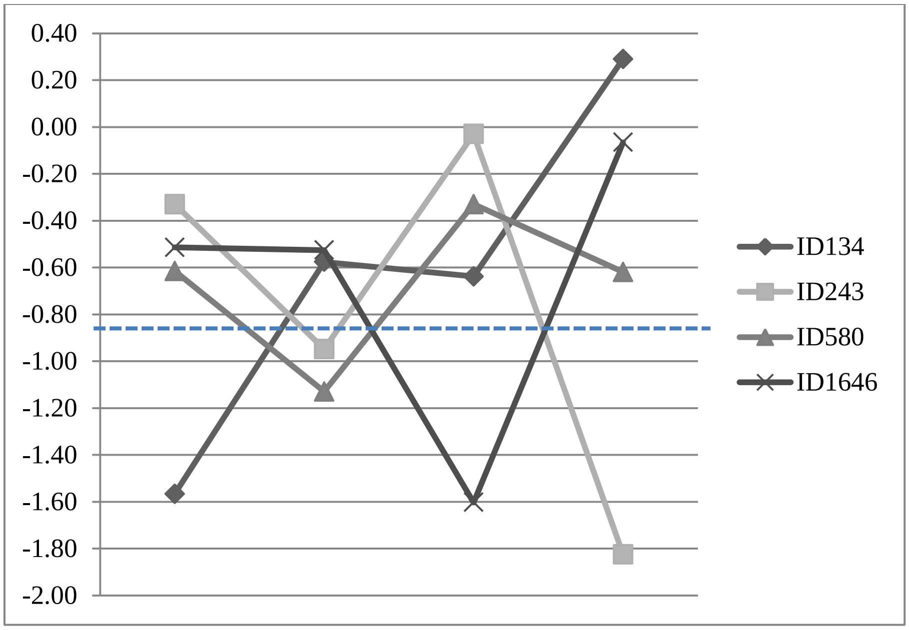

Figure 9. *MLTM-D Component Competency Patterns of Four Student Samples*

Comparison of Attribute Thetas from the MLTM-D and the CMIRT Model

The levels of component thetas from the non-compensatory MIRT model, MLTM-D, were compared with those from the CMIRT model through correlations. The Pearson correlation coefficients between thetas for the four components from the MLTM-D and the CMIRT model respectively were 0.81, 0.89, 0.86, and 0.82 in Number and Computation, Algebra, Geometry, and Data, indicating that component theta levels from those two models are very similar. Therefore, the CMIRT model seems to be able to provide diagnostic information as the MLTM-D does.

CHAPTER 5

DISCUSSION

The CMIRT models have potential for diagnostic purposes in that they can estimate person estimate for each dimension relevant to given items, and empirical reliabilities for the dimension are available. Furthermore, given dimension θ, the mastery probabilities on the dimensions as well as the probabilities of answering items correctly are obtainable. The current study showed that the CMIRT models fit the math data well, and dimension-level thetas obtained from the model showed high empirical reliabilities, indicating that the person estimates on the dimensions are reliable. Furthermore, the estimates from the CMIRT model were used to calculate dimension probabilities and pattern probabilities which enable a diagnostic interpretation.

The current study also fit the DINA model and the MLTM-D. To verify diagnostic utility of dimension probabilities from the CMIRT model, the dimension patterns were compared with those of the DINA model and the MLTM-D. In the comparison of the DINA and the CMIRT model, the number of students involved in the same pattern differed more or less depending on which cutline for the CMIRT attribute probabilities was set. The similarity in the attribute patterns of those two models was also demonstrated by high correlations between each students' dimension pattern.

The component θs from the MLTM-D were also compared with dimension θ from the CMIRT through Pearson correlations, all of which were very high, ranging from 0.82 to 0.89. From these comparisons, the diagnostic use of the CMIRT model seems to be reasonable.

The present study examined the utility of the CMIRT models in parameter estimation diagnostic potential through applications on a standardized math test administered to all middle school children in a single state for state accountability. The current study also determined the effectiveness of existing CMIRT models, and at the same time tested the relatively new model, MLTM-D, comparing the diagnostic results from both. From a psychometric perspective, when existing MIRT models are shown to have diagnostic potential, the expanded utility of the latent trait models is evident. That is, they can be a useful tool for obtaining specific and valuable information about students' cognitive abilities such as skills, knowledge, or strategies used to answer test items.

The current study has a few limitations. First, there is no direct way to compare different models based on different hypotheses and mathematical functions. Only the correlation coefficient could be used compare students' patterns across models. An additional limitation comes from the decision of which proficiency cutline level to use: the selection of an appropriate and suitable level of a cutline can offer test users highly useful information, but if too extreme a cutline is used, it could lead to wrong classification. By changing the proficiency cutline, test users are strictly or loosely classified between high ability and low ability students, meaning diverse information can be derived based on the situation in which tests are used. Finally, the present study only employed the DINA model and the MLTM-D for comparison of the competency of the dimension pattern. This may not be perfect for looking at the potential of the multidimensional latent trait models as diagnostic models. Future research should include the comparisons with additional diagnostic models.

APPENDIX A

MODEL PARAMETER ESTIMATES AND FIT INDICES

Table A1. *Classical Test Theory Statistics for 8th Grade Math Items*

Number	Item Number	Standards (Original Blueprint)	Tried	Correct	Percent Correct	Pearson Correlation	Biserial Correlation
1	p1i1	Number	2993	2867	95.8	0.21	0.47
2	p1i2	Number	2993	2615	87.4	0.30	0.48
3	p1i3	Number	2993	1704	56.9	0.28	0.36
4	p1i4	Number	2993	1861	62.2	-0.12	-0.15
5	p1i5	Number	2993	2423	81.0	0.48	0.70
6	p1i6	Number	2993	1869	62.4	0.50	0.64
7	p1i7	Number	2993	2825	94.4	0.21	0.43
8	p1i8	Number	2993	2628	87.8	0.39	0.62
9	p1i9	Number	2993	1689	56.4	0.52	0.65
10	p1i10	Number	2993	1715	57.3	0.35	0.45
11	p1i11	Number	2993	2507	83.8	0.33	0.49
12	p1i12	Number	2993	1163	38.9	0.29	0.37
13	p1i13	Data	2993	1686	56.3	0.37	0.46
14	p1i14	Data	2993	2780	92.9	0.26	0.49
15	p1i15	Data	2993	2429	81.2	0.42	0.61
16	p1i16	Data	2993	2019	67.5	0.39	0.50
17	p1i17	Data	2993	2406	80.4	0.40	0.57
18	p1i18	Data	2993	2525	84.4	0.45	0.67
19	p1i19	Geometry	2993	1910	63.8	0.49	0.63
20	p1i20	Geometry	2993	2155	72.0	0.40	0.54
21	p1i21	Geometry	2993	2129	71.1	0.56	0.74
22	p1i22	Geometry	2993	2070	69.2	0.54	0.71
23	p1i23	Geometry	2993	2347	78.4	0.47	0.66
24	p1i24	Algebra	2993	2431	81.2	0.39	0.56
25	p1i25	Algebra	2993	1999	66.8	0.48	0.62
26	p1i26	Algebra	2993	1304	43.6	0.44	0.56
27	p1i27	Algebra	2993	2257	75.4	0.37	0.50
28	p1i28	Algebra	2993	1729	57.8	0.43	0.54
29	p1i29	Algebra	2993	2249	75.1	0.50	0.68
30	p1i30	Algebra	2993	1931	64.5	0.50	0.64
31	p2i1	Algebra	2993	2629	87.8	0.30	0.49
32	p2i2	Algebra	2993	2293	76.6	0.41	0.57
33	p2i3	Algebra	2993	1740	58.1	0.36	0.45
34	p2i4	Algebra	2993	2786	93.1	0.32	0.61
35	p2i5	Algebra	2993	2261	75.5	0.49	0.67
36	p2i6	Data	2993	2287	76.4	0.26	0.36
37	p2i7	Data	2993	2773	92.6	0.37	0.69
38	p2i8	Data	2993	1618	54.1	0.24	0.30
39	p2i9	Data	2993	2553	85.3	0.44	0.68
40	p2i10	Data	2993	2497	83.4	0.45	0.68
41	p2i11	Data	2993	1952	65.2	0.44	0.57
42	p2i12	Data	2993	2578	86.1	0.43	0.67

43	p2i13	Geometry	2993	2530	84.5	0.35	0.53
44	p2i14	Geometry	2993	1904	63.6	0.53	0.68
45	p2i15	Geometry	2993	1886	63.0	0.53	0.68
46	p2i16	Geometry	2993	2367	79.1	0.45	0.63
47	p2i17	Geometry	2993	1711	57.2	0.41	0.52
48	p2i18	Number	2993	1493	49.9	0.42	0.53
49	p2i19	Number	2993	2507	83.8	0.42	0.63
50	p2i20	Number	2993	1006	33.6	0.50	0.65
51	p2i21	Number	2993	2828	94.5	0.36	0.74
52	p2i22	Number	2993	1787	59.7	0.57	0.72
53	p2i23	Number	2993	1560	52.1	0.51	0.64
54	p2i24	Data	2993	987	33.0	0.30	0.39
55	p2i25	Data	2993	1354	45.2	0.41	0.52
56	p2i26	Data	2993	1633	54.6	0.40	0.51
57	p2i27	Data	2993	1235	41.3	0.39	0.50
58	p3i1	Algebra	2993	2796	93.4	0.29	0.56
59	p3i2	Algebra	2993	2692	89.9	0.26	0.45
60	p3i3	Algebra	2993	2651	88.6	0.34	0.55
61	p3i4	Algebra	2993	2171	72.5	0.38	0.50
62	p3i5	Algebra	2993	2501	83.6	0.28	0.42
63	p3i6	Algebra	2993	1034	34.5	0.41	0.53
64	p3i7	Algebra	2993	2498	83.5	0.45	0.68
65	p3i8	Number	2993	2769	92.5	0.24	0.44
66	p3i9	Number	2993	2258	75.4	0.35	0.48
67	p3i10	Number	2993	1688	56.4	0.40	0.51
68	p3i11	Number	2993	2009	67.1	0.40	0.52
69	p3i12	Number	2993	2078	69.4	0.47	0.62
70	p3i13	Number	2993	1940	64.8	0.45	0.57
71	p3i14	Number	2993	2033	67.9	0.46	0.60
72	p3i15	Number	2993	1943	64.9	0.43	0.56
73	p3i16	Number	2993	2306	77.0	0.28	0.39
74	p3i17	Number	2993	2056	68.7	0.29	0.38
75	p3i18	Geometry	2993	2742	91.6	0.36	0.64
76	p3i19	Geometry	2993	2315	77.3	0.37	0.52
77	p3i20	Geometry	2993	2635	88.0	0.37	0.60
78	p3i21	Geometry	2993	1835	61.3	0.45	0.57
79	p3i22	Geometry	2993	1912	63.9	0.47	0.60
80	p3i23	Geometry	2993	2535	84.7	0.38	0.58
81	p3i24	Geometry	2993	2116	70.7	0.43	0.57
82	p3i25	Geometry	2993	2130	71.2	0.42	0.55
83	p3i26	Algebra	2993	2445	81.7	0.51	0.74
84	p3i27	Algebra	2993	1490	49.8	0.38	0.48
85	p3i28	Algebra	2993	2200	73.5	0.36	0.49
86	p3i29	Algebra	2993	862	28.8	0.17	0.23

Table A2. *Item Parameter Estimates from Exploratory MIRT Model*

Item	Blueprint Standard	a1	SE(a1)	a2	SE(a2)	a3	SE(a3)	a4	SE(a4)	c	SE(c)
p1i1	Number	1.38	0.15	0.41	-----	-0.29	-----	0.1	-----	4.08	0.19
p1i2	Number	1.43	0.1	-0.57	0.1	-0.56	-----	0.08	-----	2.79	0.11
p1i3	Number	0.55	0.05	-0.35	0.05	0.12	0.05	0.05	-----	0.32	0.04
p1i4	Number	-0.22	0.05	0.11	0.05	-0.02	0.06	-0.18	0.05	0.51	0.04
p1i5	Number	1.37	0.08	-0.77	0.08	0.81	0.08	0.21	0.08	2.25	0.09
p1i6	Number	1.04	0.06	-0.81	0.07	0.54	0.07	0.27	0.07	0.78	0.05
p1i7	Number	0.91	0.11	-0.59	0.12	0.32	0.11	-0.02	0.12	3.35	0.13
p1i8	Number	1.12	0.09	-0.64	0.09	0.83	0.09	0.06	0.09	2.74	0.1
p1i9	Number	1.18	0.07	-0.84	0.07	0.37	0.06	0.36	0.07	0.46	0.05
p1i10	Number	0.69	0.05	-0.45	0.06	0.22	0.06	0.08	0.06	0.37	0.04
p1i11	Number	0.67	0.07	-0.5	0.08	0.69	0.08	0.24	0.08	2.02	0.07
p1i12	Number	0.47	0.05	-0.53	0.06	0.23	0.06	0.13	0.06	-0.5	0.04
p1i13	Data	0.56	0.05	-0.54	0.06	0.48	0.06	0.2	0.06	0.33	0.04
p1i14	Data	0.8	0.1	-0.59	0.11	0.74	0.11	0.12	0.12	3.16	0.12
p1i15	Data	0.91	0.07	-0.63	0.08	0.94	0.09	0.03	0.08	2.03	0.08
p1i16	Data	0.7	0.06	-0.44	0.06	0.58	0.06	0.08	0.06	0.91	0.05
p1i17	Data	1.1	0.07	-0.54	0.08	0.39	0.07	-0.03	0.08	1.88	0.07
p1i18	Data	1.17	0.09	-0.78	0.09	1.07	0.09	0.07	0.09	2.56	0.1
p1i19	Geometry	1.2	0.07	-0.22	0.07	0.93	0.08	0.89	0.08	0.95	0.06
p1i20	Geometry	1.5	0.11	0.55	0.11	1.05	0.12	1.33	0.13	1.77	0.1
p1i21	Geometry	2.04	0.13	-0.19	0.1	1.54	0.12	1.53	0.13	2.11	0.12
p1i22	Geometry	1.41	0.08	-0.44	0.08	1.16	0.09	0.84	0.09	1.45	0.07
p1i23	Geometry	2.01	0.16	0.29	0.12	1.7	0.17	1.58	0.16	2.92	0.18
p1i24	Algebra	1.22	0.08	-0.4	0.08	0.32	0.07	0.02	0.08	1.97	0.07
p1i25	Algebra	0.86	0.06	-0.78	0.07	0.94	0.08	0.04	0.07	1.03	0.06
p1i26	Algebra	0.77	0.06	-0.61	0.06	0.58	0.07	0.24	0.06	-0.3	0.04
p1i27	Algebra	0.73	0.06	-0.51	0.07	0.6	0.07	0.07	0.07	1.4	0.06
p1i28	Algebra	0.64	0.06	-0.68	0.07	0.85	0.07	0.23	0.07	0.45	0.05

p1i29	Algebra	1.13	0.07	-0.85	0.08	0.83	0.08	0.18	0.08	1.69	0.07
p1i30	Algebra	0.96	0.06	-0.9	0.08	0.83	0.07	0.14	0.07	0.93	0.06
p2i1	Algebra	0.75	0.08	-0.36	0.09	0.65	0.09	0	0.09	2.38	0.08
p2i2	Algebra	0.85	0.06	-0.49	0.07	0.72	0.07	0.05	0.07	1.55	0.06
p2i3	Algebra	0.81	0.06	-0.43	0.06	0.09	0.06	0.04	0.06	0.41	0.04
p2i4	Algebra	1.1	0.11	-0.67	0.12	0.97	0.12	-0.01	0.13	3.57	0.14
p2i5	Algebra	1.25	0.07	-0.81	0.08	0.68	0.07	0.09	0.07	1.72	0.07
p2i6	Data	0.71	0.06	-0.25	0.07	-0.02	0.06	0.05	0.07	1.33	0.05
p2i7	Data	1.73	0.14	-0.81	0.13	0.86	0.11	0.16	0.13	4	0.18
p2i8	Data	0.37	0.05	-0.39	0.05	0.17	0.05	0.16	0.05	0.19	0.04
p2i9	Data	1.31	0.09	-0.81	0.09	0.92	0.09	0.18	0.09	2.69	0.1
p2i10	Data	1.28	0.09	-0.89	0.09	0.81	0.09	0.27	0.09	2.47	0.1
p2i11	Data	1.28	0.07	-0.59	0.07	0.05	0.07	0.05	0.07	0.92	0.05
p2i12	Data	1.31	0.09	-0.95	0.1	0.81	0.09	0.37	0.09	2.8	0.11
p2i13	Geometry	1.07	0.08	-0.35	0.08	0.41	0.07	0.17	0.08	2.17	0.08
p2i14	Geometry	1.26	0.07	-0.8	0.07	0.57	0.07	0.58	0.07	0.95	0.06
p2i15	Geometry	1.28	0.07	-0.89	0.08	0.44	0.07	0.44	0.07	0.9	0.06
p2i16	Geometry	1.24	0.08	-0.7	0.08	0.5	0.07	0.39	0.08	1.94	0.08
p2i17	Geometry	0.79	0.06	-0.49	0.06	0.41	0.06	0.28	0.06	0.39	0.04
p2i18	Number	1.05	0.06	-0.5	0.06	0.08	0.06	0.26	0.06	0.04	0.04
p2i19	Number	1.16	0.08	-0.73	0.09	0.69	0.08	0.34	0.08	2.34	0.09
p2i20	Number	1.26	0.07	-0.68	0.08	0.45	0.07	0.6	0.08	-0.95	0.05
p2i21	Number	2.07	0.17	-0.84	0.16	1.1	0.13	0.49	0.15	4.88	0.25
p2i22	Number	1.68	0.09	-0.71	0.07	0.38	0.07	0.26	0.07	0.73	0.06
p2i23	Number	1.01	0.06	-0.63	0.07	0.62	0.07	0.52	0.07	0.19	0.05
p2i24	Data	0.55	0.07	-0.86	0.08	-0.24	0.08	0.92	0.09	-0.94	0.06
p2i25	Data	0.64	0.08	-1.21	0.12	0.22	0.11	1.52	0.13	-0.23	0.06
p2i26	Data	0.9	0.07	-0.96	0.08	-0.16	0.07	0.8	0.08	0.34	0.05
p2i27	Data	0.77	0.08	-1.02	0.09	-0.07	0.09	1.18	0.1	-0.47	0.05
p3i1	Algebra	1.32	0.12	-0.36	0.12	0.51	0.11	-0.04	0.12	3.5	0.14
p3i2	Algebra	0.71	0.08	-0.38	0.09	0.58	0.09	-0.02	0.09	2.57	0.08

p3i3	Algebra	1.09	0.09	-0.33	0.09	0.58	0.09	-0.12	0.09	2.64	0.09
p3i4	Algebra	0.91	0.06	-0.44	0.07	0.31	0.06	-0.02	0.07	1.22	0.05
p3i5	Algebra	0.6	0.06	-0.35	0.07	0.47	0.07	0.03	0.07	1.86	0.06
p3i6	Algebra	1.19	0.07	-0.5	0.07	0.02	0.07	0.09	0.07	-0.82	0.05
p3i7	Algebra	1.37	0.09	-0.69	0.09	0.75	0.08	0.16	0.09	2.44	0.09
p3i8	Number	0.8	0.09	-0.17	0.1	0.62	0.1	-0.19	0.11	2.97	0.1
p3i9	Number	0.75	0.06	-0.31	0.06	0.48	0.06	0.08	0.06	1.35	0.05
p3i10	Number	1.11	0.07	-0.53	0.07	-0.1	0.06	0.06	0.07	0.38	0.05
p3i11	Number	0.83	0.06	-0.47	0.06	0.42	0.06	0.18	0.06	0.91	0.05
p3i12	Number	1.19	0.07	-0.72	0.07	0.34	0.06	0.18	0.07	1.21	0.06
p3i13	Number	0.9	0.06	-0.46	0.06	0.56	0.06	0.27	0.06	0.82	0.05
p3i14	Number	0.86	0.06	-0.79	0.08	0.65	0.07	0.37	0.07	1.08	0.06
p3i15	Number	0.7	0.06	-0.82	0.08	0.59	0.07	0.28	0.07	0.85	0.05
p3i16	Number	0.37	0.06	-0.7	0.08	0.5	0.07	0.16	0.07	1.44	0.06
p3i17	Number	0.44	0.05	-0.29	0.06	0.4	0.06	0.33	0.06	0.89	0.04
p3i18	Geometry	1.57	0.12	-0.36	0.12	0.72	0.11	0.14	0.12	3.5	0.14
p3i19	Geometry	0.87	0.06	-0.43	0.07	0.47	0.07	0.14	0.07	1.54	0.06
p3i20	Geometry	1.09	0.09	-0.71	0.09	0.68	0.09	0.26	0.09	2.71	0.1
p3i21	Geometry	0.9	0.06	-0.58	0.06	0.49	0.06	0.22	0.06	0.64	0.05
p3i22	Geometry	1.01	0.06	-0.5	0.07	0.55	0.07	0.46	0.07	0.82	0.05
p3i23	Geometry	1.2	0.08	-0.35	0.09	0.51	0.08	0.15	0.09	2.29	0.08
p3i24	Geometry	0.91	0.06	-0.43	0.07	0.62	0.07	0.37	0.07	1.18	0.05
p3i25	Geometry	0.82	0.06	-0.57	0.06	0.63	0.06	0.13	0.06	1.19	0.05
p3i26	Algebra	1.78	0.1	-0.94	0.1	0.65	0.08	0.08	0.09	2.59	0.11
p3i27	Algebra	0.77	0.05	-0.56	0.06	0.25	0.06	0.05	0.06	0.02	0.04
p3i28	Algebra	0.77	0.06	-0.48	0.06	0.35	0.06	0.28	0.06	1.26	0.05
p3i29	Algebra	0.28	0.05	-0.44	0.06	0.04	0.06	0.27	0.06	-0.97	0.04

Table A3. *Rotated Factor Loadings from Exploratory MIRT Model*

Item	**Blueprint Standard**	λ1	**SE(a1)**	λ2	**SE(a2)**	λ3	**SE(a3)**	λ4	**SE(a4)**
p1i1	Number	-0.56	0.05	0.05	0.04	-0.05	0.00	-0.33	0.03
p1i2	Number	-0.59	0.03	0.24	0.05	0.26	0.05	-0.09	0.04
p1i3	Number	-0.17	0.05	0.27	0.04	0.14	0.04	-0.11	0.04
p1i4	Number	0.07	0.06	-0.06	0.05	-0.11	0.05	0.10	0.05
p1i5	Number	-0.20	0.05	0.58	0.04	0.18	0.06	-0.33	0.05
p1i6	Number	-0.17	0.05	0.51	0.04	0.26	0.05	-0.25	0.05
p1i7	Number	-0.24	0.09	0.45	0.07	0.15	0.10	-0.15	0.09
p1i8	Number	-0.14	0.06	0.57	0.05	0.11	0.07	-0.30	0.06
p1i9	Number	-0.25	0.05	0.47	0.04	0.31	0.05	-0.26	0.05
p1i10	Number	-0.19	0.05	0.34	0.04	0.16	0.05	-0.15	0.05
p1i11	Number	-0.02	0.06	0.44	0.05	0.15	0.06	-0.28	0.06
p1i12	Number	-0.09	0.05	0.32	0.05	0.20	0.05	-0.10	0.05
p1i13	Data	-0.04	0.05	0.39	0.04	0.19	0.05	-0.20	0.05
p1i14	Data	-0.06	0.09	0.51	0.06	0.13	0.09	-0.26	0.08
p1i15	Data	-0.05	0.06	0.58	0.04	0.09	0.06	-0.28	0.05
p1i16	Data	-0.08	0.05	0.44	0.04	0.10	0.05	-0.23	0.05
p1i17	Data	-0.28	0.06	0.48	0.05	0.11	0.06	-0.21	0.05
p1i18	Data	-0.09	0.06	0.63	0.04	0.12	0.06	-0.31	0.05
p1i19	Geometry	-0.08	0.05	0.35	0.04	0.21	0.05	-0.59	0.04
p1i20	Geometry	-0.12	0.06	0.16	0.05	0.09	0.06	-0.78	0.04
p1i21	Geometry	-0.11	0.05	0.39	0.04	0.22	0.05	-0.74	0.03
p1i22	Geometry	-0.08	0.05	0.46	0.04	0.21	0.05	-0.58	0.04
p1i23	Geometry	-0.08	0.06	0.31	0.04	0.12	0.06	-0.81	0.03
p1i24	Algebra	-0.34	0.06	0.43	0.05	0.09	0.07	-0.26	0.06
p1i25	Algebra	-0.03	0.05	0.60	0.04	0.13	0.05	-0.24	0.05
p1i26	Algebra	-0.08	0.05	0.46	0.04	0.20	0.05	-0.25	0.05
p1i27	Algebra	-0.08	0.06	0.46	0.04	0.11	0.06	-0.22	0.05

p1i28	Algebra	0.05	0.05	0.51	0.04	0.18	0.05	-0.26	0.05
p1i29	Algebra	-0.13	0.05	0.59	0.04	0.20	0.05	-0.28	0.05
p1i30	Algebra	-0.07	0.05	0.60	0.04	0.21	0.05	-0.23	0.05
p2i1	Algebra	-0.08	0.07	0.45	0.06	0.03	0.08	-0.25	0.07
p2i2	Algebra	-0.09	0.06	0.50	0.04	0.08	0.06	-0.27	0.05
p2i3	Algebra	-0.28	0.05	0.33	0.04	0.15	0.05	-0.14	0.05
p2i4	Algebra	-0.10	0.08	0.61	0.06	0.08	0.09	-0.29	0.08
p2i5	Algebra	-0.21	0.05	0.58	0.04	0.18	0.05	-0.26	0.05
p2i6	Data	-0.29	0.06	0.22	0.05	0.12	0.06	-0.13	0.06
p2i7	Data	-0.28	0.07	0.61	0.05	0.16	0.08	-0.36	0.06
p2i8	Data	-0.07	0.05	0.24	0.04	0.18	0.05	-0.10	0.05
p2i9	Data	-0.16	0.06	0.61	0.04	0.17	0.06	-0.33	0.05
p2i10	Data	-0.16	0.06	0.59	0.04	0.23	0.06	-0.31	0.05
p2i11	Data	-0.42	0.05	0.40	0.04	0.20	0.06	-0.18	0.05
p2i12	Data	-0.16	0.06	0.58	0.04	0.27	0.06	-0.33	0.05
p2i13	Geometry	-0.26	0.06	0.40	0.05	0.12	0.06	-0.31	0.06
p2i14	Geometry	-0.20	0.05	0.47	0.04	0.33	0.05	-0.36	0.04
p2i15	Geometry	-0.25	0.05	0.49	0.04	0.33	0.05	-0.29	0.05
p2i16	Geometry	-0.24	0.05	0.47	0.04	0.27	0.06	-0.33	0.05
p2i17	Geometry	-0.15	0.05	0.38	0.04	0.21	0.05	-0.26	0.05
p2i18	Number	-0.33	0.05	0.33	0.04	0.24	0.05	-0.23	0.05
p2i19	Number	-0.16	0.06	0.52	0.04	0.23	0.06	-0.33	0.05
p2i20	Number	-0.23	0.05	0.42	0.05	0.32	0.05	-0.38	0.05
p2i21	Number	-0.27	0.07	0.60	0.05	0.20	0.08	-0.47	0.06
p2i22	Number	-0.39	0.05	0.49	0.04	0.22	0.05	-0.33	0.04
p2i23	Number	-0.12	0.05	0.44	0.04	0.27	0.05	-0.37	0.05
p2i24	Data	-0.15	0.07	0.13	0.05	0.58	0.05	-0.15	0.06
p2i25	Data	0.00	0.07	0.22	0.05	0.67	0.05	-0.31	0.06
p2i26	Data	-0.25	0.06	0.25	0.04	0.54	0.05	-0.18	0.05
p2i27	Data	-0.14	0.07	0.19	0.05	0.62	0.04	-0.25	0.06
p3i1	Algebra	-0.32	0.08	0.48	0.07	0.04	0.10	-0.30	0.08

p3i2	Algebra	-0.10	0.08	0.44	0.06	0.05	0.08	-0.22	0.07
p3i3	Algebra	-0.24	0.07	0.49	0.05	-0.01	0.08	-0.27	0.07
p3i4	Algebra	-0.25	0.05	0.41	0.04	0.10	0.06	-0.19	0.05
p3i5	Algebra	-0.09	0.06	0.38	0.05	0.07	0.07	-0.19	0.06
p3i6	Algebra	-0.40	0.05	0.36	0.05	0.19	0.06	-0.19	0.05
p3i7	Algebra	-0.23	0.05	0.57	0.04	0.16	0.06	-0.33	0.05
p3i8	Number	-0.14	0.08	0.43	0.07	-0.09	0.09	-0.24	0.08
p3i9	Number	-0.14	0.06	0.38	0.05	0.07	0.06	-0.25	0.05
p3i10	Number	-0.42	0.05	0.33	0.04	0.21	0.05	-0.14	0.05
p3i11	Number	-0.17	0.05	0.40	0.04	0.17	0.05	-0.25	0.05
p3i12	Number	-0.29	0.05	0.48	0.04	0.24	0.05	-0.23	0.05
p3i13	Number	-0.14	0.05	0.43	0.04	0.17	0.05	-0.31	0.05
p3i14	Number	-0.07	0.05	0.49	0.04	0.28	0.05	-0.27	0.05
p3i15	Number	-0.04	0.06	0.49	0.05	0.28	0.05	-0.20	0.05
p3i16	Number	0.05	0.07	0.41	0.05	0.23	0.06	-0.11	0.06
p3i17	Number	-0.01	0.06	0.26	0.05	0.17	0.05	-0.25	0.05
p3i18	Geometry	-0.31	0.07	0.51	0.06	0.06	0.09	-0.41	0.07
p3i19	Geometry	-0.17	0.05	0.42	0.05	0.13	0.06	-0.26	0.05
p3i20	Geometry	-0.16	0.06	0.52	0.05	0.21	0.07	-0.30	0.06
p3i21	Geometry	-0.16	0.05	0.45	0.04	0.20	0.05	-0.26	0.05
p3i22	Geometry	-0.16	0.05	0.41	0.04	0.23	0.05	-0.37	0.05
p3i23	Geometry	-0.27	0.06	0.44	0.05	0.10	0.07	-0.34	0.06
p3i24	Geometry	-0.12	0.06	0.41	0.05	0.18	0.06	-0.36	0.05
p3i25	Geometry	-0.10	0.05	0.49	0.04	0.15	0.05	-0.25	0.05
p3i26	Algebra	-0.34	0.05	0.61	0.04	0.20	0.06	-0.29	0.05
p3i27	Algebra	-0.21	0.05	0.40	0.04	0.18	0.05	-0.14	0.05
p3i28	Algebra	-0.16	0.05	0.36	0.04	0.21	0.05	-0.25	0.05
p3i29	Algebra	-0.06	0.06	0.17	0.05	0.26	0.05	-0.07	0.06

Table A4. *Common Item Discrimination, and Difficulty Parameter Estimates of 1PL model*

Item	Blueprint Standard	Item Discrimination(a)	Standard Error(a)	Item Difficulty(b)	Standard Error(b)
p1i1	Number	1.16	0.02	-3.16	0.09
p1i2	Number	1.16	0.02	-2.04	0.06
p1i3	Number	1.16	0.02	-0.32	0.04
p1i4	Number	1.16	0.02	-0.56	0.04
p1i5	Number	1.16	0.02	-1.56	0.05
p1i6	Number	1.16	0.02	-0.57	0.04
p1i7	Number	1.16	0.02	-2.88	0.08
p1i8	Number	1.16	0.02	-2.08	0.06
p1i9	Number	1.16	0.02	-0.30	0.04
p1i10	Number	1.16	0.02	-0.34	0.04
p1i11	Number	1.16	0.02	-1.76	0.05
p1i12	Number	1.16	0.02	0.48	0.04
p1i13	Data	1.16	0.02	-0.30	0.04
p1i14	Data	1.16	0.02	-2.65	0.08
p1i15	Data	1.16	0.02	-1.57	0.05
p1i16	Data	1.16	0.02	-0.81	0.04
p1i17	Data	1.16	0.02	-1.52	0.05
p1i18	Data	1.16	0.02	-1.80	0.05
p1i19	Geometry	1.16	0.02	-0.64	0.04
p1i20	Geometry	1.16	0.02	-1.04	0.04
p1i21	Geometry	1.16	0.02	-1.00	0.04
p1i22	Geometry	1.16	0.02	-0.90	0.04
p1i23	Geometry	1.16	0.02	-1.40	0.05
p1i24	Algebra	1.16	0.02	-1.58	0.05
p1i25	Algebra	1.16	0.02	-0.78	0.04
p1i26	Algebra	1.16	0.02	0.27	0.04
p1i27	Algebra	1.16	0.02	-1.23	0.05
p1i28	Algebra	1.16	0.02	-0.36	0.04
p1i29	Algebra	1.16	0.02	-1.21	0.04
p1i30	Algebra	1.16	0.02	-0.67	0.04
p2i1	Algebra	1.16	0.02	-2.09	0.06
p2i2	Algebra	1.16	0.02	-1.29	0.05
p2i3	Algebra	1.16	0.02	-0.38	0.04
p2i4	Algebra	1.16	0.02	-2.68	0.08
p2i5	Algebra	1.16	0.02	-1.23	0.05
p2i6	Data	1.16	0.02	-1.28	0.05
p2i7	Data	1.16	0.02	-2.62	0.07
p2i8	Data	1.16	0.02	-0.20	0.04
p2i9	Data	1.16	0.02	-1.87	0.06
p2i10	Data	1.16	0.02	-1.73	0.05
p2i11	Data	1.16	0.02	-0.70	0.04
p2i12	Data	1.16	0.02	-1.94	0.06

p2i13	Geometry	1.16	0.02	-1.81	0.05
p2i14	Geometry	1.16	0.02	-0.63	0.04
p2i15	Geometry	1.16	0.02	-0.60	0.04
p2i16	Geometry	1.16	0.02	-1.44	0.05
p2i17	Geometry	1.16	0.02	-0.33	0.04
p2i18	Number	1.16	0.02	-0.01	0.04
p2i19	Number	1.16	0.02	-1.76	0.05
p2i20	Number	1.16	0.02	0.74	0.04
p2i21	Number	1.16	0.02	-2.90	0.08
p2i22	Number	1.16	0.02	-0.45	0.04
p2i23	Number	1.16	0.02	-0.11	0.04
p2i24	Data	1.16	0.02	0.77	0.04
p2i25	Data	1.16	0.02	0.19	0.04
p2i26	Data	1.16	0.02	-0.22	0.04
p2i27	Data	1.16	0.02	0.37	0.04
p3i1	Algebra	1.16	0.02	-2.73	0.08
p3i2	Algebra	1.16	0.02	-2.29	0.07
p3i3	Algebra	1.16	0.02	-2.15	0.06
p3i4	Algebra	1.16	0.02	-1.07	0.04
p3i5	Algebra	1.16	0.02	-1.74	0.05
p3i6	Algebra	1.16	0.02	0.69	0.04
p3i7	Algebra	1.16	0.02	-1.73	0.05
p3i8	Number	1.16	0.02	-2.60	0.07
p3i9	Number	1.16	0.02	-1.23	0.05
p3i10	Number	1.16	0.02	-0.30	0.04
p3i11	Number	1.16	0.02	-0.80	0.04
p3i12	Number	1.16	0.02	-0.91	0.04
p3i13	Number	1.16	0.02	-0.69	0.04
p3i14	Number	1.16	0.02	-0.83	0.04
p3i15	Number	1.16	0.02	-0.69	0.04
p3i16	Number	1.16	0.02	-1.32	0.05
p3i17	Number	1.16	0.02	-0.87	0.04
p3i18	Geometry	1.16	0.02	-2.48	0.07
p3i19	Geometry	1.16	0.02	-1.34	0.05
p3i20	Geometry	1.16	0.02	-2.10	0.06
p3i21	Geometry	1.16	0.02	-0.52	0.04
p3i22	Geometry	1.16	0.02	-0.64	0.04
p3i23	Geometry	1.16	0.02	-1.83	0.05
p3i24	Geometry	1.16	0.02	-0.97	0.04
p3i25	Geometry	1.16	0.02	-1.00	0.04
p3i26	Algebra	1.16	0.02	-1.61	0.05
p3i27	Algebra	1.16	0.02	-0.01	0.04
p3i28	Algebra	1.16	0.02	-1.12	0.04
p3i29	Algebra	1.16	0.02	0.98	0.04

Table A5. *Item Discrimination and Difficulty Parameter Estimates from 2PL Model*

Item	Blueprint Standard	Item Discrimination(a)	Standard Error(a)	Item Difficulty(b)	Standard Error(b)
p1i1	Number	1.28	0.14	-2.93	0.24
p1i2	Number	1.14	0.08	-2.06	0.11
p1i3	Number	0.67	0.04	-0.47	0.07
p1i4	Number	0.01	0.01	-75.52	59.74
p1i5	Number	1.87	0.09	-1.21	0.04
p1i6	Number	1.44	0.06	-0.53	0.04
p1i7	Number	1.13	0.11	-2.93	0.23
p1i8	Number	1.63	0.10	-1.69	0.07
p1i9	Number	1.53	0.07	-0.29	0.04
p1i10	Number	0.86	0.05	-0.42	0.05
p1i11	Number	1.08	0.07	-1.84	0.10
p1i12	Number	0.72	0.04	0.69	0.07
p1i13	Data	0.91	0.05	-0.35	0.05
p1i14	Data	1.26	0.11	-2.49	0.16
p1i15	Data	1.42	0.08	-1.40	0.06
p1i16	Data	1.01	0.05	-0.90	0.06
p1i17	Data	1.31	0.07	-1.42	0.06
p1i18	Data	1.79	0.10	-1.41	0.05
p1i19	Geometry	1.48	0.07	-0.58	0.04
p1i20	Geometry	1.19	0.06	-1.03	0.05
p1i21	Geometry	2.03	0.09	-0.78	0.03
p1i22	Geometry	1.84	0.08	-0.73	0.03
p1i23	Geometry	1.73	0.09	-1.13	0.04
p1i24	Algebra	1.32	0.07	-1.46	0.06
p1i25	Algebra	1.37	0.06	-0.72	0.04
p1i26	Algebra	1.16	0.05	0.26	0.05
p1i27	Algebra	1.07	0.06	-1.29	0.07
p1i28	Algebra	1.14	0.05	-0.38	0.04
p1i29	Algebra	1.67	0.08	-1.00	0.04
p1i30	Algebra	1.49	0.07	-0.60	0.04
p2i1	Algebra	1.07	0.08	-2.20	0.13
p2i2	Algebra	1.22	0.07	-1.26	0.06
p2i3	Algebra	0.87	0.05	-0.46	0.05
p2i4	Algebra	1.72	0.14	-2.08	0.10
p2i5	Algebra	1.68	0.08	-1.02	0.04
p2i6	Data	0.68	0.05	-1.91	0.14
p2i7	Data	2.26	0.16	-1.79	0.06
p2i8	Data	0.57	0.04	-0.32	0.07
p2i9	Data	1.88	0.10	-1.43	0.05
p2i10	Data	1.83	0.10	-1.35	0.05
p2i11	Data	1.24	0.06	-0.69	0.04
p2i12	Data	1.86	0.10	-1.49	0.05
p2i13	Geometry	1.25	0.08	-1.73	0.08

p2i14	Geometry	1.69	0.07	-0.54	0.03
p2i15	Geometry	1.64	0.07	-0.52	0.03
p2i16	Geometry	1.58	0.08	-1.21	0.05
p2i17	Geometry	1.06	0.05	-0.36	0.04
p2i18	Number	1.12	0.05	-0.02	0.04
p2i19	Number	1.62	0.09	-1.44	0.05
p2i20	Number	1.62	0.07	0.59	0.04
p2i21	Number	2.82	0.21	-1.80	0.06
p2i22	Number	1.80	0.08	-0.39	0.03
p2i23	Number	1.44	0.06	-0.12	0.04
p2i24	Data	0.83	0.05	0.97	0.07
p2i25	Data	1.09	0.05	0.19	0.05
p2i26	Data	1.07	0.05	-0.24	0.04
p2i27	Data	1.07	0.05	0.38	0.05
p3i1	Algebra	1.53	0.13	-2.26	0.12
p3i2	Algebra	1.02	0.09	-2.50	0.17
p3i3	Algebra	1.32	0.09	-1.97	0.10
p3i4	Algebra	1.05	0.06	-1.14	0.06
p3i5	Algebra	0.85	0.07	-2.17	0.14
p3i6	Algebra	1.14	0.05	0.69	0.05
p3i7	Algebra	1.80	0.10	-1.36	0.05
p3i8	Number	1.01	0.10	-2.87	0.22
p3i9	Number	0.96	0.06	-1.39	0.08
p3i10	Number	1.02	0.05	-0.33	0.04
p3i11	Number	1.09	0.06	-0.84	0.05
p3i12	Number	1.43	0.07	-0.82	0.04
p3i13	Number	1.22	0.06	-0.68	0.04
p3i14	Number	1.35	0.07	-0.78	0.04
p3i15	Number	1.18	0.06	-0.69	0.04
p3i16	Number	0.79	0.06	-1.73	0.11
p3i17	Number	0.71	0.05	-1.25	0.09
p3i18	Geometry	1.87	0.13	-1.85	0.07
p3i19	Geometry	1.13	0.07	-1.36	0.06
p3i20	Geometry	1.55	0.10	-1.76	0.07
p3i21	Geometry	1.23	0.06	-0.52	0.04
p3i22	Geometry	1.34	0.06	-0.61	0.04
p3i23	Geometry	1.41	0.09	-1.61	0.07
p3i24	Geometry	1.25	0.06	-0.94	0.05
p3i25	Geometry	1.20	0.06	-0.98	0.05
p3i26	Algebra	2.16	0.11	-1.18	0.04
p3i27	Algebra	0.97	0.05	-0.01	0.05
p3i28	Algebra	1.03	0.06	-1.22	0.06
p3i29	Algebra	0.51	0.04	1.89	0.17

Table A6. *Item Discrimination, Difficulty, and Guessing Parameter Estimates of 3PL Model*

Item	Blueprint Standard	Item Discrimination(a)	Standard Error(a)	Item Difficulty(b)	Standard Error(b)	Item Guessing(g)	Standard Error(g)
p1i1	Number	1.10	0.11	-3.05	0.29	0.26	0.09
p1i2	Number	1.39	0.17	-0.98	0.25	0.52	0.07
p1i3	Number	1.18	0.16	0.60	0.13	0.32	0.04
p1i4	Number	0.01	0.01	14.13	40.42	0.29	0.12
p1i5	Number	1.82	0.13	-1.01	0.10	0.17	0.05
p1i6	Number	2.05	0.15	-0.05	0.06	0.21	0.03
p1i7	Number	0.96	0.10	-2.97	0.34	0.29	0.10
p1i8	Number	1.52	0.12	-1.52	0.17	0.24	0.08
p1i9	Number	2.83	0.20	0.21	0.04	0.22	0.02
p1i10	Number	1.40	0.15	0.37	0.10	0.27	0.03
p1i11	Number	1.07	0.09	-1.50	0.25	0.23	0.09
p1i12	Number	1.68	0.18	1.11	0.06	0.22	0.02
p1i13	Data	1.58	0.15	0.41	0.08	0.28	0.03
p1i14	Data	1.10	0.10	-2.41	0.27	0.29	0.10
p1i15	Data	1.36	0.10	-1.24	0.15	0.14	0.07
p1i16	Data	1.16	0.11	-0.42	0.16	0.19	0.06
p1i17	Data	1.32	0.11	-1.08	0.18	0.22	0.07
p1i18	Data	1.70	0.13	-1.24	0.13	0.19	0.07
p1i19	Geometry	2.28	0.17	-0.02	0.05	0.25	0.03
p1i20	Geometry	1.53	0.14	-0.36	0.12	0.30	0.05
p1i21	Geometry	2.41	0.17	-0.43	0.06	0.19	0.03
p1i22	Geometry	2.45	0.17	-0.30	0.05	0.22	0.03
p1i23	Geometry	2.03	0.16	-0.67	0.09	0.28	0.04
p1i24	Algebra	1.59	0.15	-0.76	0.15	0.37	0.06
p1i25	Algebra	1.54	0.12	-0.40	0.09	0.15	0.04
p1i26	Algebra	1.63	0.13	0.53	0.05	0.12	0.02
p1i27	Algebra	1.35	0.14	-0.51	0.16	0.33	0.06

p1i28	Algebra	1.80	0.15	0.20	0.07	0.23	0.03
p1i29	Algebra	1.92	0.14	-0.60	0.09	0.23	0.04
p1i30	Algebra	2.17	0.16	-0.09	0.06	0.24	0.03
p2i1	Algebra	1.03	0.09	-1.94	0.27	0.23	0.10
p2i2	Algebra	1.22	0.10	-1.04	0.17	0.13	0.07
p2i3	Algebra	1.17	0.12	0.16	0.13	0.21	0.05
p2i4	Algebra	1.47	0.12	-2.10	0.21	0.25	0.10
p2i5	Algebra	1.88	0.14	-0.65	0.09	0.21	0.04
p2i6	Data	0.73	0.08	-1.32	0.42	0.21	0.11
p2i7	Data	1.98	0.19	-1.67	0.17	0.33	0.08
p2i8	Data	1.61	0.24	1.00	0.08	0.39	0.02
p2i9	Data	1.79	0.14	-1.24	0.12	0.21	0.06
p2i10	Data	2.02	0.16	-0.93	0.11	0.30	0.05
p2i11	Data	1.99	0.17	0.02	0.07	0.30	0.03
p2i12	Data	1.89	0.16	-1.14	0.13	0.30	0.06
p2i13	Geometry	1.22	0.11	-1.40	0.23	0.25	0.09
p2i14	Geometry	2.86	0.21	-0.01	0.04	0.25	0.02
p2i15	Geometry	2.31	0.16	-0.09	0.05	0.20	0.02
p2i16	Geometry	2.01	0.17	-0.60	0.10	0.34	0.04
p2i17	Geometry	1.53	0.13	0.19	0.08	0.22	0.03
p2i18	Number	2.09	0.17	0.50	0.05	0.22	0.02
p2i19	Number	2.08	0.20	-0.72	0.11	0.43	0.05
p2i20	Number	2.76	0.19	0.72	0.03	0.08	0.01
p2i21	Number	2.36	0.23	-1.75	0.15	0.35	0.09
p2i22	Number	2.10	0.13	-0.14	0.05	0.10	0.02
p2i23	Number	1.91	0.13	0.18	0.05	0.12	0.02
p2i24	Data	3.67	0.38	1.10	0.04	0.20	0.01
p2i25	Data	2.90	0.30	0.72	0.04	0.24	0.02
p2i26	Data	3.00	0.32	0.56	0.04	0.33	0.02
p2i27	Data	3.79	0.37	0.80	0.03	0.23	0.01
p3i1	Algebra	1.30	0.11	-2.35	0.23	0.22	0.11

p3i2	Algebra	0.96	0.09	-2.30	0.31	0.23	0.11
p3i3	Algebra	1.22	0.10	-1.90	0.25	0.17	0.11
p3i4	Algebra	1.40	0.14	-0.34	0.15	0.33	0.05
p3i5	Algebra	0.86	0.08	-1.76	0.35	0.22	0.11
p3i6	Algebra	2.11	0.17	0.88	0.04	0.13	0.01
p3i7	Algebra	1.87	0.15	-1.02	0.11	0.27	0.05
p3i8	Number	0.96	0.09	-2.73	0.28	0.22	0.09
p3i9	Number	1.00	0.09	-1.04	0.23	0.17	0.08
p3i10	Number	1.62	0.15	0.30	0.07	0.24	0.03
p3i11	Number	1.56	0.14	-0.09	0.10	0.30	0.04
p3i12	Number	2.02	0.16	-0.22	0.07	0.28	0.03
p3i13	Number	1.36	0.11	-0.35	0.11	0.14	0.04
p3i14	Number	2.19	0.19	-0.05	0.07	0.32	0.03
p3i15	Number	1.75	0.16	-0.02	0.09	0.28	0.03
p3i16	Number	1.13	0.17	-0.35	0.27	0.45	0.07
p3i17	Number	0.85	0.11	-0.45	0.30	0.25	0.08
p3i18	Geometry	1.68	0.16	-1.68	0.20	0.32	0.09
p3i19	Geometry	1.20	0.10	-0.94	0.18	0.22	0.07
p3i20	Geometry	1.58	0.15	-1.30	0.18	0.37	0.08
p3i21	Geometry	1.78	0.15	0.05	0.08	0.24	0.03
p3i22	Geometry	1.99	0.16	-0.03	0.07	0.25	0.03
p3i23	Geometry	1.37	0.12	-1.33	0.18	0.24	0.08
p3i24	Geometry	1.53	0.13	-0.41	0.11	0.24	0.05
p3i25	Geometry	1.44	0.12	-0.46	0.12	0.24	0.05
p3i26	Algebra	2.53	0.19	-0.78	0.07	0.29	0.04
p3i27	Algebra	1.82	0.17	0.58	0.06	0.24	0.02
p3i28	Algebra	1.44	0.15	-0.30	0.15	0.37	0.05
p3i29	Algebra	2.74	0.35	1.53	0.06	0.21	0.01

Table A7. *Item Discrimination, Difficulty, and Guessing Parameter Estimates of 3PL Model with Common Asymptote*

Item	**Blueprint Standard**	**Item Discrimination(a)**	**Standard Error(a)**	**Item Difficulty(b)**	**Standard Error(b)**	**Item Guessing(g)**	**Standard Error(g)**
p1i1	Number	1.13	0.11	-3.06	0.24	0.19	0.01
p1i2	Number	1.05	0.07	-1.93	0.12	0.19	0.01
p1i3	Number	0.85	0.06	0.18	0.07	0.19	0.01
p1i4	Number	0.01	0.01	-16.76	14.80	0.19	0.01
p1i5	Number	1.88	0.10	-1.00	0.05	0.19	0.01
p1i6	Number	1.92	0.11	-0.12	0.04	0.19	0.01
p1i7	Number	0.99	0.09	-3.05	0.25	0.19	0.01
p1i8	Number	1.52	0.09	-1.58	0.08	0.19	0.01
p1i9	Number	2.51	0.15	0.14	0.04	0.19	0.01
p1i10	Number	1.13	0.07	0.13	0.05	0.19	0.01
p1i11	Number	1.06	0.07	-1.59	0.10	0.19	0.01
p1i12	Number	1.44	0.12	1.09	0.07	0.19	0.01
p1i13	Data	1.24	0.08	0.18	0.05	0.19	0.01
p1i14	Data	1.10	0.09	-2.54	0.17	0.19	0.01
p1i15	Data	1.41	0.08	-1.16	0.07	0.19	0.01
p1i16	Data	1.14	0.07	-0.45	0.06	0.19	0.01
p1i17	Data	1.31	0.08	-1.15	0.07	0.19	0.01
p1i18	Data	1.73	0.10	-1.24	0.06	0.19	0.01
p1i19	Geometry	1.98	0.11	-0.16	0.04	0.19	0.01
p1i20	Geometry	1.32	0.08	-0.64	0.06	0.19	0.01
p1i21	Geometry	2.38	0.13	-0.46	0.04	0.19	0.01
p1i22	Geometry	2.32	0.13	-0.38	0.04	0.19	0.01
p1i23	Geometry	1.84	0.10	-0.86	0.05	0.19	0.01
p1i24	Algebra	1.33	0.08	-1.19	0.07	0.19	0.01
p1i25	Algebra	1.60	0.09	-0.34	0.04	0.19	0.01
p1i26	Algebra	1.82	0.13	0.66	0.05	0.19	0.01
p1i27	Algebra	1.14	0.07	-0.91	0.06	0.19	0.01

p1i28	Algebra	1.60	0.09	0.09	0.04	0.19	0.01
p1i29	Algebra	1.83	0.10	-0.69	0.05	0.19	0.01
p1i30	Algebra	1.97	0.11	-0.19	0.04	0.19	0.01
p2i1	Algebra	1.03	0.08	-2.01	0.12	0.19	0.01
p2i2	Algebra	1.26	0.07	-0.94	0.06	0.19	0.01
p2i3	Algebra	1.10	0.07	0.08	0.05	0.19	0.01
p2i4	Algebra	1.50	0.11	-2.13	0.12	0.19	0.01
p2i5	Algebra	1.83	0.10	-0.72	0.05	0.19	0.01
p2i6	Data	0.72	0.06	-1.38	0.11	0.19	0.01
p2i7	Data	1.94	0.14	-1.83	0.09	0.19	0.01
p2i8	Data	0.76	0.06	0.42	0.07	0.19	0.01
p2i9	Data	1.81	0.11	-1.28	0.06	0.19	0.01
p2i10	Data	1.84	0.11	-1.14	0.05	0.19	0.01
p2i11	Data	1.57	0.09	-0.24	0.04	0.19	0.01
p2i12	Data	1.78	0.11	-1.34	0.06	0.19	0.01
p2i13	Geometry	1.21	0.08	-1.52	0.08	0.19	0.01
p2i14	Geometry	2.46	0.14	-0.13	0.04	0.19	0.01
p2i15	Geometry	2.20	0.12	-0.14	0.04	0.19	0.01
p2i16	Geometry	1.66	0.10	-0.93	0.05	0.19	0.01
p2i17	Geometry	1.42	0.09	0.12	0.05	0.19	0.01
p2i18	Number	1.84	0.13	0.43	0.05	0.19	0.01
p2i19	Number	1.59	0.09	-1.23	0.06	0.19	0.01
p2i20	Number	3.42	0.28	0.85	0.05	0.19	0.01
p2i21	Number	2.31	0.19	-1.90	0.09	0.19	0.01
p2i22	Number	2.31	0.13	-0.04	0.04	0.19	0.01
p2i23	Number	2.04	0.13	0.27	0.04	0.19	0.01
p2i24	Data	3.11	0.30	1.12	0.05	0.19	0.01
p2i25	Data	2.13	0.18	0.62	0.05	0.19	0.01
p2i26	Data	1.70	0.12	0.26	0.04	0.19	0.01
p2i27	Data	2.91	0.26	0.76	0.04	0.19	0.01
p3i1	Algebra	1.36	0.11	-2.30	0.14	0.19	0.01

p3i2	Algebra	0.97	0.08	-2.36	0.16	0.19	0.01
p3i3	Algebra	1.26	0.09	-1.84	0.10	0.19	0.01
p3i4	Algebra	1.16	0.07	-0.72	0.06	0.19	0.01
p3i5	Algebra	0.86	0.06	-1.83	0.12	0.19	0.01
p3i6	Algebra	2.41	0.21	0.99	0.05	0.19	0.01
p3i7	Algebra	1.78	0.10	-1.16	0.05	0.19	0.01
p3i8	Number	0.97	0.08	-2.74	0.20	0.19	0.01
p3i9	Number	1.01	0.07	-1.00	0.07	0.19	0.01
p3i10	Number	1.40	0.09	0.16	0.05	0.19	0.01
p3i11	Number	1.28	0.08	-0.38	0.05	0.19	0.01
p3i12	Number	1.72	0.10	-0.43	0.04	0.19	0.01
p3i13	Number	1.42	0.08	-0.27	0.05	0.19	0.01
p3i14	Number	1.64	0.09	-0.37	0.04	0.19	0.01
p3i15	Number	1.46	0.09	-0.25	0.05	0.19	0.01
p3i16	Number	0.82	0.06	-1.29	0.10	0.19	0.01
p3i17	Number	0.79	0.06	-0.67	0.08	0.19	0.01
p3i18	Geometry	1.65	0.12	-1.86	0.10	0.19	0.01
p3i19	Geometry	1.18	0.07	-1.03	0.06	0.19	0.01
p3i20	Geometry	1.44	0.10	-1.64	0.09	0.19	0.01
p3i21	Geometry	1.58	0.10	-0.07	0.04	0.19	0.01
p3i22	Geometry	1.74	0.10	-0.18	0.04	0.19	0.01
p3i23	Geometry	1.36	0.08	-1.43	0.07	0.19	0.01
p3i24	Geometry	1.42	0.08	-0.55	0.05	0.19	0.01
p3i25	Geometry	1.35	0.08	-0.59	0.05	0.19	0.01
p3i26	Algebra	2.28	0.15	-0.95	0.04	0.19	0.01
p3i27	Algebra	1.54	0.11	0.48	0.05	0.19	0.01
p3i28	Algebra	1.12	0.07	-0.80	0.06	0.19	0.01
p3i29	Algebra	2.25	0.26	1.57	0.08	0.19	0.01

Table A8. *S-X^2 Item-level Fit from 1PL Model*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	51.50	50	0.4160
p1i2	Number	84.85	58	0.0123
p1i3	Number	185.06	59	0.0001
p1i4	Number	2013.99	59	0.0001
p1i5	Number	103.03	59	0.0003
p1i6	Number	92.73	59	0.0033
p1i7	Number	68.18	54	0.0927
p1i8	Number	69.15	57	0.1296
p1i9	Number	142.02	59	0.0001
p1i10	Number	103.04	58	0.0002
p1i11	Number	63.17	58	0.2981
p1i12	Number	208.87	56	0.0001
p1i13	Data	138.29	59	0.0001
p1i14	Data	55.36	55	0.4621
p1i15	Data	80.42	59	0.0333
p1i16	Data	79.31	59	0.0400
p1i17	Data	49.18	59	0.8156
p1i18	Data	87.91	58	0.0068
p1i19	Geometry	97.80	59	0.0011
p1i20	Geometry	79.25	60	0.0486
p1i21	Geometry	131.63	59	0.0001
p1i22	Geometry	141.61	59	0.0001
p1i23	Geometry	77.30	60	0.0656
p1i24	Algebra	65.37	59	0.2646
p1i25	Algebra	53.44	59	0.6804
p1i26	Algebra	61.12	58	0.3639
p1i27	Algebra	66.95	60	0.2504
p1i28	Algebra	110.66	59	0.0001
p1i29	Algebra	98.12	59	0.0010
p1i30	Algebra	100.90	60	0.0007
p2i1	Algebra	94.12	58	0.0019
p2i2	Algebra	61.59	59	0.3828
p2i3	Algebra	111.56	58	0.0001
p2i4	Algebra	72.40	55	0.0577
p2i5	Algebra	86.99	59	0.0103
p2i6	Data	164.36	60	0.0001
p2i7	Data	101.93	55	0.0001
p2i8	Data	291.10	59	0.0001
p2i9	Data	106.61	58	0.0001
p2i10	Data	107.68	59	0.0001
p2i11	Data	95.91	60	0.0022
p2i12	Data	97.71	58	0.0009
p2i13	Geometry	35.42	58	0.9916

p2i14	Geometry	153.88	60	0.0001
p2i15	Geometry	104.84	59	0.0002
p2i16	Geometry	68.45	59	0.1869
p2i17	Geometry	87.11	59	0.0101
p2i18	Number	93.70	58	0.0021
p2i19	Number	72.00	59	0.1189
p2i20	Number	165.57	55	0.0001
p2i21	Number	108.70	54	0.0001
p2i22	Number	126.77	59	0.0001
p2i23	Number	69.68	58	0.1397
p2i24	Data	322.26	54	0.0001
p2i25	Data	110.09	58	0.0001
p2i26	Data	140.97	58	0.0001
p2i27	Data	213.40	58	0.0001
p3i1	Algebra	73.84	54	0.0376
p3i2	Algebra	59.48	57	0.3846
p3i3	Algebra	75.12	58	0.0646
p3i4	Algebra	60.46	59	0.4240
p3i5	Algebra	90.01	58	0.0045
p3i6	Algebra	115.56	56	0.0001
p3i7	Algebra	85.26	59	0.0142
p3i8	Number	80.18	56	0.0187
p3i9	Number	77.29	59	0.0552
p3i10	Number	68.30	58	0.1667
p3i11	Number	56.43	58	0.5349
p3i12	Number	56.05	59	0.5857
p3i13	Number	50.19	60	0.8130
p3i14	Number	76.36	58	0.0534
p3i15	Number	64.10	60	0.3341
p3i16	Number	119.75	59	0.0001
p3i17	Number	151.84	59	0.0001
p3i18	Geometry	73.73	56	0.0562
p3i19	Geometry	51.58	60	0.7727
p3i20	Geometry	63.31	58	0.2940
p3i21	Geometry	67.03	60	0.2482
p3i22	Geometry	64.79	60	0.3126
p3i23	Geometry	43.29	58	0.9250
p3i24	Geometry	34.43	59	0.9956
p3i25	Geometry	50.34	59	0.7820
p3i26	Algebra	148.46	59	0.0001
p3i27	Algebra	118.72	58	0.0001
p3i28	Algebra	72.97	59	0.1042
p3i29	Algebra	709.14	55	0.0001
45 out of 86 items fitted well at the p-level of 0.01				

Table A9. *S-X^2 Item-level Fit from 2PL Model*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	61.13	49	0.1143
p1i2	Number	80.99	58	0.0247
p1i3	Number	42.32	61	0.9673
p1i4	Number	135.91	62	0.0001
p1i5	Number	56.12	54	0.3946
p1i6	Number	77.68	58	0.0431
p1i7	Number	66.45	54	0.1188
p1i8	Number	51.20	53	0.5454
p1i9	Number	118.23	57	0.0001
p1i10	Number	48.84	60	0.8484
p1i11	Number	63.63	59	0.3163
p1i12	Number	64.24	60	0.3298
p1i13	Data	82.12	60	0.0305
p1i14	Data	57.74	54	0.3380
p1i15	Data	88.66	57	0.0046
p1i16	Data	67.83	60	0.2273
p1i17	Data	47.28	57	0.8175
p1i18	Data	58.35	54	0.3179
p1i19	Geometry	82.72	58	0.0182
p1i20	Geometry	78.46	60	0.0550
p1i21	Geometry	44.00	53	0.8063
p1i22	Geometry	86.00	55	0.0047
p1i23	Geometry	49.52	56	0.7174
p1i24	Algebra	64.08	58	0.2713
p1i25	Algebra	50.24	59	0.7851
p1i26	Algebra	59.82	58	0.4106
p1i27	Algebra	59.22	60	0.5051
p1i28	Algebra	97.36	59	0.0012
p1i29	Algebra	64.28	57	0.2364
p1i30	Algebra	82.37	58	0.0194
p2i1	Algebra	93.86	58	0.0020
p2i2	Algebra	69.65	59	0.1615
p2i3	Algebra	65.52	60	0.2908
p2i4	Algebra	59.33	50	0.1715
p2i5	Algebra	58.01	57	0.4389
p2i6	Data	63.53	61	0.3868
p2i7	Data	52.71	45	0.2002
p2i8	Data	69.29	62	0.2445
p2i9	Data	75.07	53	0.0247
p2i10	Data	69.28	54	0.0786
p2i11	Data	88.15	59	0.0083
p2i12	Data	66.04	54	0.1258
p2i13	Geometry	37.15	58	0.9850

p2i14	Geometry	125.55	57	0.0001
p2i15	Geometry	62.94	56	0.2436
p2i16	Geometry	49.77	56	0.7087
p2i17	Geometry	74.70	59	0.0815
p2i18	Number	80.06	59	0.0353
p2i19	Number	49.09	56	0.7322
p2i20	Number	152.43	52	0.0001
p2i21	Number	27.98	39	0.9054
p2i22	Number	57.78	56	0.4104
p2i23	Number	54.95	58	0.5903
p2i24	Data	182.92	59	0.0001
p2i25	Data	88.39	58	0.0062
p2i26	Data	113.68	58	0.0001
p2i27	Data	169.68	58	0.0001
p3i1	Algebra	75.32	52	0.0189
p3i2	Algebra	57.35	58	0.5004
p3i3	Algebra	75.86	57	0.0480
p3i4	Algebra	52.66	60	0.7387
p3i5	Algebra	55.55	60	0.6394
p3i6	Algebra	98.68	57	0.0005
p3i7	Algebra	56.59	54	0.3778
p3i8	Number	78.61	56	0.0248
p3i9	Number	62.24	60	0.3956
p3i10	Number	53.95	60	0.6959
p3i11	Number	48.70	59	0.8285
p3i12	Number	40.02	58	0.9656
p3i13	Number	52.01	59	0.7290
p3i14	Number	66.56	58	0.2058
p3i15	Number	58.28	59	0.5031
p3i16	Number	62.42	60	0.3894
p3i17	Number	53.43	61	0.7443
p3i18	Geometry	42.88	49	0.7190
p3i19	Geometry	51.19	60	0.7842
p3i20	Geometry	56.63	55	0.4150
p3i21	Geometry	61.26	60	0.4315
p3i22	Geometry	56.01	59	0.5871
p3i23	Geometry	41.03	57	0.9453
p3i24	Geometry	33.04	59	0.9975
p3i25	Geometry	50.16	59	0.7873
p3i26	Algebra	69.25	52	0.0549
p3i27	Algebra	80.89	59	0.0308
p3i28	Algebra	59.37	60	0.4995
p3i29	Algebra	185.04	60	0.0001
71 out of 86 items fitted well at the p-level of 0.01				

Table A10. *S-X^2 Item-level Fit from 3PL Model*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	57.26	49	0.1949
p1i2	Number	79.86	57	0.0245
p1i3	Number	39.41	60	0.9816
p1i4	Number	135.74	61	0.0001
p1i5	Number	46.35	55	0.7908
p1i6	Number	58.26	57	0.4299
p1i7	Number	66.43	55	0.1387
p1i8	Number	46.93	55	0.7727
p1i9	Number	46.05	57	0.8502
p1i10	Number	41.75	60	0.9649
p1i11	Number	52.49	59	0.7131
p1i12	Number	43.68	60	0.9441
p1i13	Data	67.98	60	0.2236
p1i14	Data	56.91	56	0.4421
p1i15	Data	66.15	57	0.1899
p1i16	Data	61.69	59	0.3794
p1i17	Data	48.24	57	0.7896
p1i18	Data	45.92	55	0.8040
p1i19	Geometry	60.24	58	0.3939
p1i20	Geometry	76.96	59	0.0581
p1i21	Geometry	50.08	55	0.6635
p1i22	Geometry	67.63	55	0.1177
p1i23	Geometry	46.05	56	0.8263
p1i24	Algebra	63.07	57	0.2698
p1i25	Algebra	43.22	58	0.9261
p1i26	Algebra	55.66	58	0.5638
p1i27	Algebra	58.20	59	0.5061
p1i28	Algebra	73.17	58	0.0864
p1i29	Algebra	59.57	56	0.3463
p1i30	Algebra	61.85	57	0.3066
p2i1	Algebra	74.11	58	0.0753
p2i2	Algebra	52.92	58	0.6648
p2i3	Algebra	66.06	59	0.2458
p2i4	Algebra	47.61	51	0.6098
p2i5	Algebra	51.44	56	0.6487
p2i6	Data	54.39	60	0.6805
p2i7	Data	53.53	46	0.2072
p2i8	Data	52.23	60	0.7524
p2i9	Data	66.48	53	0.1008
p2i10	Data	63.09	54	0.1855
p2i11	Data	72.64	58	0.0932
p2i12	Data	54.50	54	0.4563
p2i13	Geometry	37.67	58	0.9824

p2i14	Geometry	67.80	57	0.1547
p2i15	Geometry	52.40	57	0.6486
p2i16	Geometry	48.06	55	0.7355
p2i17	Geometry	67.25	59	0.2151
p2i18	Number	48.48	59	0.8343
p2i19	Number	44.54	55	0.8425
p2i20	Number	60.78	56	0.3072
p2i21	Number	27.00	39	0.9269
p2i22	Number	55.42	55	0.4598
p2i23	Number	45.71	56	0.8355
p2i24	Data	61.01	59	0.4029
p2i25	Data	56.03	58	0.5496
p2i26	Data	66.93	57	0.1725
p2i27	Data	49.85	58	0.7684
p3i1	Algebra	53.37	52	0.4227
p3i2	Algebra	47.41	57	0.8137
p3i3	Algebra	60.67	57	0.3444
p3i4	Algebra	51.77	59	0.7371
p3i5	Algebra	44.77	59	0.9149
p3i6	Algebra	56.39	57	0.4988
p3i7	Algebra	50.30	55	0.6550
p3i8	Number	55.46	57	0.5340
p3i9	Number	50.40	59	0.7803
p3i10	Number	44.34	59	0.9221
p3i11	Number	45.84	58	0.8763
p3i12	Number	35.51	57	0.9886
p3i13	Number	49.17	59	0.8158
p3i14	Number	57.96	57	0.4408
p3i15	Number	52.34	58	0.6854
p3i16	Number	64.72	59	0.2834
p3i17	Number	51.78	60	0.7666
p3i18	Geometry	41.28	50	0.8058
p3i19	Geometry	48.45	59	0.8351
p3i20	Geometry	54.85	55	0.4815
p3i21	Geometry	52.46	58	0.6813
p3i22	Geometry	39.81	59	0.9740
p3i23	Geometry	35.33	57	0.9893
p3i24	Geometry	33.45	58	0.9960
p3i25	Geometry	49.17	59	0.8158
p3i26	Algebra	64.47	51	0.0972
p3i27	Algebra	52.34	59	0.7182
p3i28	Algebra	56.89	59	0.5545
p3i29	Algebra	87.65	60	0.0114
85 out of 86 items fitted well at the p-level of 0.01				

Table A11. *S-X^2 Item-level Fit from 3PLC Model*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	47.01	49	0.5549
p1i2	Number	83.54	58	0.0157
p1i3	Number	43.93	60	0.9408
p1i4	Number	135.84	61	0.0001
p1i5	Number	50.18	54	0.6231
p1i6	Number	60.12	57	0.3626
p1i7	Number	66.42	55	0.1389
p1i8	Number	46.91	53	0.7095
p1i9	Number	50.56	57	0.7142
p1i10	Number	45.62	59	0.8993
p1i11	Number	54.78	59	0.6324
p1i12	Number	45.25	60	0.9215
p1i13	Data	75.56	60	0.0847
p1i14	Data	57.58	55	0.3791
p1i15	Data	73.97	56	0.0541
p1i16	Data	63.49	59	0.3206
p1i17	Data	48.22	57	0.7900
p1i18	Data	47.03	54	0.7384
p1i19	Geometry	64.56	57	0.2290
p1i20	Geometry	77.84	59	0.0506
p1i21	Geometry	48.36	54	0.6914
p1i22	Geometry	69.09	55	0.0957
p1i23	Geometry	46.58	55	0.7839
p1i24	Algebra	65.82	57	0.1976
p1i25	Algebra	49.75	58	0.7716
p1i26	Algebra	72.56	58	0.0943
p1i27	Algebra	59.39	59	0.4622
p1i28	Algebra	79.16	59	0.0410
p1i29	Algebra	59.26	56	0.3570
p1i30	Algebra	65.57	57	0.2037
p2i1	Algebra	78.63	57	0.0303
p2i2	Algebra	60.46	58	0.3862
p2i3	Algebra	65.87	59	0.2510
p2i4	Algebra	50.68	51	0.4874
p2i5	Algebra	53.12	56	0.5855
p2i6	Data	56.56	60	0.6029
p2i7	Data	53.88	46	0.1980
p2i8	Data	72.06	61	0.1569
p2i9	Data	69.11	53	0.0676
p2i10	Data	67.33	53	0.0889
p2i11	Data	81.47	58	0.0227
p2i12	Data	56.58	54	0.3783
p2i13	Geometry	37.75	58	0.9819

p2i14	Geometry	75.27	56	0.0438
p2i15	Geometry	52.75	56	0.5995
p2i16	Geometry	49.26	56	0.7265
p2i17	Geometry	68.83	59	0.1785
p2i18	Number	51.07	59	0.7595
p2i19	Number	55.63	56	0.4899
p2i20	Number	120.52	58	0.0001
p2i21	Number	29.03	39	0.8782
p2i22	Number	72.81	57	0.0771
p2i23	Number	62.75	58	0.3109
p2i24	Data	61.94	59	0.3711
p2i25	Data	58.23	59	0.5048
p2i26	Data	97.79	59	0.0011
p2i27	Data	56.74	58	0.5233
p3i1	Algebra	67.36	52	0.0743
p3i2	Algebra	49.39	57	0.7534
p3i3	Algebra	58.57	56	0.3805
p3i4	Algebra	53.92	59	0.6635
p3i5	Algebra	46.98	59	0.8710
p3i6	Algebra	73.58	59	0.0957
p3i7	Algebra	51.27	55	0.6185
p3i8	Number	60.16	57	0.3613
p3i9	Number	54.60	60	0.6734
p3i10	Number	46.64	59	0.8785
p3i11	Number	50.89	59	0.7653
p3i12	Number	37.82	57	0.9765
p3i13	Number	51.89	58	0.7012
p3i14	Number	66.83	58	0.1994
p3i15	Number	55.67	58	0.5633
p3i16	Number	68.17	59	0.1932
p3i17	Number	51.76	60	0.7672
p3i18	Geometry	43.25	50	0.7395
p3i19	Geometry	48.95	59	0.8219
p3i20	Geometry	54.88	56	0.5184
p3i21	Geometry	54.47	58	0.6082
p3i22	Geometry	43.72	58	0.9179
p3i23	Geometry	36.54	57	0.9841
p3i24	Geometry	33.31	58	0.9962
p3i25	Geometry	49.67	59	0.8019
p3i26	Algebra	68.14	52	0.0657
p3i27	Algebra	57.78	59	0.5215
p3i28	Algebra	61.35	59	0.3910
p3i29	Algebra	96.40	60	0.0020
82 out of 86 items fitted well at the p-level of 0.01				

Table A12. *Item Factor Loadings from Full Information Bifactor Model*

Item	Blueprint Standard	λ1	SE(g)	λ2	SE(d1)	λ3	SE(d2)	λ4	SE(d3)	λ5	SE(d4)
p1i1	Number	**0.60**	0.70	**-0.60**	0.11						
p1i2	Number	**0.57**	0.50	**-0.80**	0.70						
p1i3	Number	**0.37**	0.40	**0.70**	0.50						
p1i4	Number	-0.15	0.40	**-0.70**	0.60						
p1i5	Number	**0.74**	0.30	**0.30**	0.50						
p1i6	Number	**0.65**	0.30	0.10	0.50						
p1i7	Number	**0.56**	0.70	**0.80**	0.10						
p1i8	Number	**0.69**	0.40	**-0.40**	0.70						
p1i9	Number	**0.67**	0.30	-0.10	0.50						
p1i10	Number	**0.46**	0.30	-0.10	0.50						
p1i11	Number	**0.54**	0.40	-0.10	0.70						
p1i12	Number	**0.40**	0.30	**0.40**	0.60						
p1i13	Data	**0.48**	0.30							0.10	0.50
p1i14	Data	**0.59**	0.60							**-0.70**	0.10
p1i15	Data	**0.63**	0.40							-0.10	0.70
p1i16	Data	**0.51**	0.30							**-0.80**	0.60
p1i17	Data	**0.61**	0.40							**-0.50**	0.60
p1i18	Data	**0.72**	0.30							-0.10	0.70
p1i19	Geometry	**0.63**	0.30					**0.37**	0.40		
p1i20	Geometry	**0.54**	0.30					**0.62**	0.50		
p1i21	Geometry	**0.73**	0.20					**0.49**	0.40		
p1i22	Geometry	**0.71**	0.30					**0.31**	0.40		
p1i23	Geometry	**0.68**	0.30					**0.56**	0.40		
p1i24	Algebra	**0.62**	0.40			0.10	0.70				
p1i25	Algebra	**0.62**	0.30			**0.32**	0.60				
p1i26	Algebra	**0.56**	0.30			0.26	0.60				
p1i27	Algebra	**0.53**	0.40			0.24	0.60				

p1i28	Algebra	**0.54**	0.30			**0.53**	0.60				
p1i29	Algebra	**0.69**	0.30			0.28	0.60				
p1i30	Algebra	**0.65**	0.30			**0.42**	0.60				
p2i1	Algebra	**0.54**	0.50			0.10	0.80				
p2i2	Algebra	**0.59**	0.40				0.60				
p2i3	Algebra	**0.47**	0.30			**-0.60**	0.60				
p2i4	Algebra	**0.71**	0.50			**-0.70**	0.90				
p2i5	Algebra	**0.70**	0.30			0.11	0.60				
p2i6	Data	**0.38**	0.40							**0.30**	0.60
p2i7	Data	**0.80**	0.30							-0.10	0.80
p2i8	Data	**0.32**	0.40							**0.60**	0.50
p2i9	Data	**0.74**	0.30							-0.10	0.70
p2i10	Data	**0.74**	0.30							**0.30**	0.60
p2i11	Data	**0.60**	0.30							-0.20	0.50
p2i12	Data	**0.75**	0.30							**0.60**	0.60
p2i13	Geometry	**0.59**	0.40					**0.60**	0.60		
p2i14	Geometry	**0.70**	0.30					**0.70**	0.50		
p2i15	Geometry	**0.70**	0.30					**-0.30**	0.50		
p2i16	Geometry	**0.68**	0.30					0.20	0.50		
p2i17	Geometry	**0.53**	0.30					**0.40**	0.50		
p2i18	Number	**0.56**	0.30	**-0.40**	0.50						
p2i19	Number	**0.69**	0.30	**0.60**	0.60						
p2i20	Number	**0.69**	0.30	**0.30**	0.50						
p2i21	Number	**0.85**	0.30	**0.30**	0.70						
p2i22	Number	**0.74**	0.20	**-0.70**	0.50						
p2i23	Number	**0.65**	0.30	**0.80**	0.50						
p2i24	Data	**0.42**	0.40							**0.52**	0.50
p2i25	Data	**0.54**	0.30							**0.62**	0.50
p2i26	Data	**0.54**	0.30							**0.39**	0.50
p2i27	Data	**0.52**	0.30							**0.53**	0.50
p3i1	Algebra	**0.67**	0.50			-0.12	0.10				

p3i2	Algebra	**0.52**	0.50			0.10	0.90		
p3i3	Algebra	**0.62**	0.50			-0.10	0.80		
p3i4	Algebra	**0.53**	0.40				0.60		
p3i5	Algebra	**0.45**	0.50			**0.80**	0.80		
p3i6	Algebra	**0.57**	0.30			**-0.30**	0.60		
p3i7	Algebra	**0.73**	0.30			0.20	0.60		
p3i8	Number	**0.52**	0.60	-0.10	0.90				
p3i9	Number	**0.50**	0.40		0.60				
p3i10	Number	**0.53**	0.30	-0.10	0.50				
p3i11	Number	**0.54**	0.30	**0.30**	0.50				
p3i12	Number	**0.65**	0.30	0.10	0.50				
p3i13	Number	**0.58**	0.30	**0.40**	0.50				
p3i14	Number	**0.62**	0.30	**0.34**	0.50				
p3i15	Number	**0.57**	0.30	**0.67**	0.80				
p3i16	Number	**0.42**	0.40	**0.55**	0.70				
p3i17	Number	**0.38**	0.40	0.21	0.60				
p3i18	Geometry	**0.74**	0.40					**0.90**	0.70
p3i19	Geometry	**0.56**	0.40					0.10	0.60
p3i20	Geometry	**0.67**	0.40					0.10	0.60
p3i21	Geometry	**0.59**	0.30					0.10	0.50
p3i22	Geometry	**0.61**	0.30					0.10	0.50
p3i23	Geometry	**0.64**	0.40					**0.50**	0.60
p3i24	Geometry	**0.59**	0.30					0.10	0.60
p3i25	Geometry	**0.58**	0.30					**0.30**	0.50
p3i26	Algebra	**0.79**	0.30			**0.70**	0.60		
p3i27	Algebra	**0.50**	0.30			**0.80**	0.60		
p3i28	Algebra	**0.52**	0.40			0.10	0.60		
p3i29	Algebra	0.28	0.40			**0.70**	0.70		

Table A13. *Parameter Estimates from Full Information Bifactor Model*

Item	Blueprint Standard	g	SE(g)	a1	SE(a1)	a2	SE(a2)	a3	SE(a3)	a4	SE(a4)	c	SE(c)
p1i1	Number	1.28	0.14	-0.13	0.14	0.00	-----	0.00	-----	0.00	-----	3.75	0.15
p1i2	Number	1.17	0.08	-0.17	0.09	0.00	-----	0.00	-----	0.00	-----	2.38	0.09
p1i3	Number	0.68	0.04	0.12	0.06	0.00	-----	0.00	-----	0.00	-----	0.32	0.04
p1i4	Number	-0.27	0.04	-0.13	0.06	0.00	-----	0.00	-----	0.00	-----	0.51	0.04
p1i5	Number	1.89	0.09	0.06	0.08	0.00	-----	0.00	-----	0.00	-----	2.27	0.10
p1i6	Number	1.46	0.07	0.02	0.07	0.00	-----	0.00	-----	0.00	-----	0.76	0.06
p1i7	Number	1.14	0.11	0.16	0.12	0.00	-----	0.00	-----	0.00	-----	3.32	0.12
p1i8	Number	1.64	0.10	-0.10	0.09	0.00	-----	0.00	-----	0.00	-----	2.76	0.11
p1i9	Number	1.55	0.07	-0.03	0.07	0.00	-----	0.00	-----	0.00	-----	0.43	0.06
p1i10	Number	0.88	0.05	-0.02	0.06	0.00	-----	0.00	-----	0.00	-----	0.36	0.05
p1i11	Number	1.09	0.07	-0.02	0.08	0.00	-----	0.00	-----	0.00	-----	1.99	0.07
p1i12	Number	0.73	0.05	0.07	0.06	0.00	-----	0.00	-----	0.00	-----	-0.50	0.04
p1i13	Data	0.92	0.05	0.00	-----	0.00	-----	0.00	-----	0.02	0.06	0.32	0.05
p1i14	Data	1.25	0.11	0.00	-----	0.00	-----	0.00	-----	-0.14	0.12	3.14	0.12
p1i15	Data	1.41	0.08	0.00	-----	0.00	-----	0.00	-----	-0.23	0.09	1.98	0.08
p1i16	Data	1.02	0.06	0.00	-----	0.00	-----	0.00	-----	-0.15	0.07	0.90	0.05
p1i17	Data	1.32	0.07	0.00	-----	0.00	-----	0.00	-----	-0.11	0.08	1.86	0.07
p1i18	Data	1.78	0.10	0.00	-----	0.00	-----	0.00	-----	-0.25	0.10	2.53	0.10
p1i19	Geometry	1.59	0.08	0.00	-----	0.00	-----	0.93	0.08	0.00	-----	0.95	0.07
p1i20	Geometry	1.59	0.11	0.00	-----	0.00	-----	1.82	0.16	0.00	-----	1.79	0.12
p1i21	Geometry	2.65	0.15	0.00	-----	0.00	-----	1.78	0.15	0.00	-----	2.18	0.15
p1i22	Geometry	1.93	0.09	0.00	-----	0.00	-----	0.84	0.08	0.00	-----	1.45	0.08
p1i23	Geometry	2.41	0.16	0.00	-----	0.00	-----	2.00	0.18	0.00	-----	2.90	0.18
p1i24	Algebra	1.33	0.08	0.00	-----	0.02	0.08	0.00	-----	0.00	-----	1.92	0.07
p1i25	Algebra	1.47	0.07	0.00	-----	0.77	0.09	0.00	-----	0.00	-----	1.08	0.07
p1i26	Algebra	1.19	0.06	0.00	-----	0.56	0.08	0.00	-----	0.00	-----	-0.32	0.05
p1i27	Algebra	1.10	0.06	0.00	-----	0.51	0.08	0.00	-----	0.00	-----	1.45	0.06
p1i28	Algebra	1.40	0.09	0.00	-----	1.36	0.16	0.00	-----	0.00	-----	0.55	0.07

p1i29	Algebra	1.77	0.09	0.00	-----	0.72	0.10	0.00	-----	0.00	-----	1.81	0.09
p1i30	Algebra	1.71	0.09	0.00	-----	1.10	0.12	0.00	-----	0.00	-----	1.07	0.08
p2i1	Algebra	1.08	0.08	0.00	-----	0.02	0.10	0.00	-----	0.00	-----	2.36	0.08
p2i2	Algebra	1.23	0.07	0.00	-----	0.00	0.08	0.00	-----	0.00	-----	1.54	0.06
p2i3	Algebra	0.91	0.05	0.00	-----	-0.12	0.07	0.00	-----	0.00	-----	0.40	0.05
p2i4	Algebra	1.74	0.13	0.00	-----	-0.18	0.13	0.00	-----	0.00	-----	3.59	0.15
p2i5	Algebra	1.70	0.08	0.00	-----	0.26	0.08	0.00	-----	0.00	-----	1.72	0.08
p2i6	Data	0.70	0.05	0.00	-----	0.00	-----	0.00	-----	0.05	0.07	1.30	0.05
p2i7	Data	2.28	0.16	0.00	-----	0.00	-----	0.00	-----	-0.02	0.13	4.05	0.19
p2i8	Data	0.58	0.04	0.00	-----	0.00	-----	0.00	-----	0.12	0.06	0.18	0.04
p2i9	Data	1.90	0.11	0.00	-----	0.00	-----	0.00	-----	-0.02	0.10	2.70	0.11
p2i10	Data	1.85	0.10	0.00	-----	0.00	-----	0.00	-----	0.07	0.09	2.47	0.10
p2i11	Data	1.27	0.06	0.00	-----	0.00	-----	0.00	-----	-0.05	0.07	0.86	0.06
p2i12	Data	1.91	0.11	0.00	-----	0.00	-----	0.00	-----	0.16	0.10	2.80	0.12
p2i13	Geometry	1.25	0.08	0.00	-----	0.00	-----	0.13	0.08	0.00	-----	2.16	0.08
p2i14	Geometry	1.69	0.07	0.00	-----	0.00	-----	0.18	0.07	0.00	-----	0.91	0.06
p2i15	Geometry	1.68	0.07	0.00	-----	0.00	-----	-0.07	0.07	0.00	-----	0.86	0.06
p2i16	Geometry	1.59	0.08	0.00	-----	0.00	-----	0.05	0.07	0.00	-----	1.92	0.08
p2i17	Geometry	1.07	0.05	0.00	-----	0.00	-----	0.07	0.06	0.00	-----	0.38	0.05
p2i18	Number	1.15	0.05	-0.08	0.06	0.00	-----	0.00	-----	0.00	-----	0.02	0.05
p2i19	Number	1.63	0.09	0.15	0.08	0.00	-----	0.00	-----	0.00	-----	2.35	0.09
p2i20	Number	1.64	0.07	0.06	0.07	0.00	-----	0.00	-----	0.00	-----	-0.96	0.06
p2i21	Number	2.77	0.21	0.11	0.14	0.00	-----	0.00	-----	0.00	-----	5.00	0.28
p2i22	Number	1.87	0.08	-0.17	0.07	0.00	-----	0.00	-----	0.00	-----	0.71	0.07
p2i23	Number	1.45	0.06	0.18	0.07	0.00	-----	0.00	-----	0.00	-----	0.17	0.06
p2i24	Data	0.95	0.06	0.00	-----	0.00	-----	0.00	-----	1.20	0.10	-0.97	0.06
p2i25	Data	1.59	0.12	0.00	-----	0.00	-----	0.00	-----	1.82	0.19	-0.25	0.07
p2i26	Data	1.24	0.07	0.00	-----	0.00	-----	0.00	-----	0.89	0.08	0.32	0.05
p2i27	Data	1.33	0.08	0.00	-----	0.00	-----	0.00	-----	1.34	0.12	-0.49	0.06
p3i1	Algebra	1.55	0.13	0.00	-----	-0.28	0.14	0.00	-----	0.00	-----	3.50	0.15
p3i2	Algebra	1.03	0.09	0.00	-----	0.01	0.10	0.00	-----	0.00	-----	2.56	0.09

p3i3	Algebra	1.36	0.10	0.00	-----	-0.21	0.11	0.00	-----	0.00	-----	2.64	0.10
p3i4	Algebra	1.07	0.06	0.00	-----	0.00	0.07	0.00	-----	0.00	-----	1.21	0.06
p3i5	Algebra	0.85	0.07	0.00	-----	0.15	0.09	0.00	-----	0.00	-----	1.85	0.06
p3i6	Algebra	1.18	0.06	0.00	-----	-0.05	0.07	0.00	-----	0.00	-----	-0.79	0.05
p3i7	Algebra	1.82	0.10	0.00	-----	0.04	0.09	0.00	-----	0.00	-----	2.45	0.10
p3i8	Number	1.03	0.10	-0.01	0.11	0.00	-----	0.00	-----	0.00	-----	2.90	0.10
p3i9	Number	0.97	0.06	0.01	0.07	0.00	-----	0.00	-----	0.00	-----	1.34	0.06
p3i10	Number	1.05	0.05	-0.03	0.06	0.00	-----	0.00	-----	0.00	-----	0.34	0.05
p3i11	Number	1.10	0.06	0.07	0.06	0.00	-----	0.00	-----	0.00	-----	0.91	0.05
p3i12	Number	1.46	0.07	0.03	0.07	0.00	-----	0.00	-----	0.00	-----	1.18	0.06
p3i13	Number	1.22	0.06	0.09	0.06	0.00	-----	0.00	-----	0.00	-----	0.82	0.05
p3i14	Number	1.49	0.08	0.83	0.09	0.00	-----	0.00	-----	0.00	-----	1.17	0.07
p3i15	Number	2.07	0.28	2.44	0.47	0.00	-----	0.00	-----	0.00	-----	1.48	0.20
p3i16	Number	0.98	0.08	1.30	0.15	0.00	-----	0.00	-----	0.00	-----	1.76	0.10
p3i17	Number	0.71	0.05	0.39	0.07	0.00	-----	0.00	-----	0.00	-----	0.90	0.05
p3i18	Geometry	1.89	0.13	0.00	-----	0.00	-----	0.23	0.11	0.00	-----	3.50	0.15
p3i19	Geometry	1.14	0.07	0.00	-----	0.00	-----	0.03	0.07	0.00	-----	1.54	0.06
p3i20	Geometry	1.55	0.10	0.00	-----	0.00	-----	0.03	0.09	0.00	-----	2.71	0.11
p3i21	Geometry	1.24	0.06	0.00	-----	0.00	-----	0.03	0.06	0.00	-----	0.63	0.05
p3i22	Geometry	1.34	0.06	0.00	-----	0.00	-----	0.21	0.07	0.00	-----	0.81	0.06
p3i23	Geometry	1.42	0.09	0.00	-----	0.00	-----	0.12	0.08	0.00	-----	2.29	0.09
p3i24	Geometry	1.25	0.06	0.00	-----	0.00	-----	0.21	0.07	0.00	-----	1.17	0.06
p3i25	Geometry	1.21	0.06	0.00	-----	0.00	-----	0.05	0.06	0.00	-----	1.18	0.06
p3i26	Algebra	2.19	0.11	0.00	-----	0.21	0.09	0.00	-----	0.00	-----	2.58	0.11
p3i27	Algebra	0.99	0.05	0.00	-----	0.15	0.07	0.00	-----	0.00	-----	0.01	0.05
p3i28	Algebra	1.04	0.06	0.00	-----	0.03	0.07	0.00	-----	0.00	-----	1.25	0.06
p3i29	Algebra	0.50	0.04	0.00	-----	0.12	0.07	0.00	-----	0.00	-----	-0.96	0.04

Table A14. *Parameter Estimates from Original Blueprint based CMIRT Model*

Item	Blueprint Standard	a1	SE(a1)	a2	SE(a2)	a3	SE(a3)	a4	SE(a4)	c	SE(c)
p1i1	Number	0.88	0.11							3.54	0.13
p1i2	Number	0.98	0.07							2.34	0.08
p1i3	Number	0.54	0.04							0.30	0.04
p1i4	Number	-0.28	0.04							0.51	0.04
p1i5	Number	**1.40**	0.07							2.09	0.08
p1i6	Number	**1.18**	0.06							0.69	0.05
p1i7	Number	0.85	0.09							3.21	0.11
p1i8	Number	**1.22**	0.08							2.60	0.09
p1i9	Number	**1.24**	0.06							0.36	0.05
p1i10	Number	0.73	0.04							0.34	0.04
p1i11	Number	0.85	0.06							1.93	0.07
p1i12	Number	0.60	0.04							-0.51	0.04
p1i13	Data							0.80	0.05	0.25	0.04
p1i14	Data							**1.25**	0.10	3.14	0.12
p1i15	Data							**1.25**	0.07	1.83	0.07
p1i16	Data							0.90	0.06	0.81	0.05
p1i17	Data							**1.13**	0.07	1.71	0.06
p1i18	Data							**1.56**	0.09	2.32	0.09
p1i19	Geometry					**1.29**	0.06			0.86	0.05
p1i20	Geometry					**1.21**	0.06			1.35	0.06
p1i21	Geometry					**1.73**	0.07			1.58	0.07
p1i22	Geometry					**1.53**	0.07			1.32	0.06
p1i23	Geometry					**1.64**	0.08			2.14	0.09
p1i24	Algebra			**1.09**	0.06					1.90	0.07
p1i25	Algebra			**1.27**	0.06					1.01	0.05
p1i26	Algebra			**1.07**	0.06					-0.30	0.04
p1i27	Algebra			**0.96**	0.06					1.41	0.06

p1i28	Algebra			**1.09**	0.05					0.44	0.05
p1i29	Algebra			**1.43**	0.07					1.66	0.07
p1i30	Algebra			**1.37**	0.06					0.91	0.05
p2i1	Algebra			0.95	0.07					2.39	0.08
p2i2	Algebra			0.98	0.06					1.50	0.06
p2i3	Algebra			0.76	0.05					0.40	0.04
p2i4	Algebra			1.22	0.10					3.33	0.13
p2i5	Algebra			1.41	0.07					1.68	0.07
p2i6	Data							0.62	0.05	1.24	0.05
p2i7	Data							**1.89**	0.14	3.75	0.17
p2i8	Data							0.48	0.04	0.14	0.04
p2i9	Data							**1.57**	0.09	2.43	0.09
p2i10	Data							**1.59**	0.08	2.25	0.08
p2i11	Data							0.94	0.06	0.70	0.04
p2i12	Data							**1.50**	0.09	2.46	0.09
p2i13	Geometry					0.93	0.06			2.08	0.07
p2i14	Geometry					**1.30**	0.06			0.85	0.05
p2i15	Geometry					**1.21**	0.06			0.78	0.05
p2i16	Geometry					**1.16**	0.06			1.81	0.07
p2i17	Geometry					0.89	0.05			0.39	0.04
p2i18	Number	0.92	0.05							-0.01	0.04
p2i19	Number	**1.26**	0.07							2.22	0.08
p2i20	Number	**1.40**	0.07							-1.01	0.06
p2i21	Number	**2.08**	0.20							4.66	0.28
p2i22	Number	**1.40**	0.06							0.58	0.05
p2i23	Number	**1.22**	0.06							0.12	0.05
p2i24	Data							**1.00**	0.06	-0.92	0.05
p2i25	Data							**1.39**	0.08	-0.33	0.05
p2i26	Data							**1.14**	0.06	0.17	0.04
p2i27	Data							**1.23**	0.07	-0.53	0.05
p3i1	Algebra			**1.19**	0.10					3.36	0.13

p3i2	Algebra			0.86	0.07					2.55	0.08
p3i3	Algebra			**1.04**	0.07					2.54	0.09
p3i4	Algebra			0.89	0.05					1.20	0.05
p3i5	Algebra			0.73	0.06					1.85	0.06
p3i6	Algebra			**1.01**	0.06					-0.78	0.05
p3i7	Algebra			**1.39**	0.08					2.33	0.09
p3i8	Number	0.76	0.08							2.81	0.09
p3i9	Number	0.81	0.05							1.31	0.05
p3i10	Number	0.85	0.05							0.31	0.04
p3i11	Number	0.90	0.05							0.87	0.05
p3i12	Number	**1.23**	0.06							1.13	0.06
p3i13	Number	**1.00**	0.05							0.78	0.05
p3i14	Number	**1.12**	0.05							1.00	0.05
p3i15	Number	**1.04**	0.05							0.79	0.05
p3i16	Number	0.73	0.05							1.38	0.05
p3i17	Number	0.59	0.04							0.86	0.04
p3i18	Geometry					**1.46**	0.11			3.40	0.14
p3i19	Geometry					0.91	0.05			1.52	0.06
p3i20	Geometry					**1.10**	0.08			2.56	0.09
p3i21	Geometry					0.96	0.05			0.62	0.05
p3i22	Geometry					**1.17**	0.05			0.82	0.05
p3i23	Geometry					**1.14**	0.07			2.26	0.08
p3i24	Geometry					**1.13**	0.06			1.22	0.06
p3i25	Geometry					0.93	0.05			1.15	0.05
p3i26	Algebra			**1.59**	0.08					2.33	0.09
p3i27	Algebra			0.81	0.05					0.01	0.04
p3i28	Algebra			0.81	0.05					1.22	0.05
p3i29	Algebra			0.37	0.04					-0.93	0.04

Table A15. *Parameter Estimates from Augmented Blueprint based CMIRT Model*

Item	Blueprint Standard	a1	SE(a1)	a2	SE(a2)	a3	SE(a3)	a4	SE(a4)	c	SE(c)
p1i1	Number	0.38	0.13	0.81	0.13					3.70	0.15
p1i2	Number	0.59	0.08	0.51	0.08					2.31	0.08
p1i3	Number	0.30	0.05	0.38	0.05					0.31	0.04
p1i4	Number	-0.18	0.05	-0.11	0.05					0.51	0.04
p1i5	Number	0.70	0.08	1.08	0.08					2.20	0.08
p1i6	Number	1.24	0.06							0.68	0.05
p1i7	Number	0.50	0.12	0.62	0.11					3.29	0.12
p1i8	Number	1.19	0.08							2.52	0.09
p1i9	Number	1.25	0.06							0.35	0.05
p1i10	Number	0.78	0.05							0.34	0.04
p1i11	Number	0.94	0.07							1.95	0.07
p1i12	Number	0.66	0.05							-0.51	0.04
p1i13	Data	0.54	0.06	0.25	0.06			0.21	0.07	0.29	0.04
p1i14	Data	0.78	0.11					0.52	0.11	3.07	0.11
p1i15	Data	0.64	0.08	0.58	0.08			0.26	0.09	1.92	0.07
p1i16	Data	0.43	0.06	0.49	0.06			0.17	0.07	0.88	0.05
p1i17	Data	0.50	0.08	0.55	0.07			0.38	0.08	1.81	0.07
p1i18	Data	0.83	0.09	0.74	0.09			0.32	0.10	2.47	0.09
p1i19	Geometry					1.29	0.06			0.87	0.05
p1i20	Geometry					1.21	0.06			1.35	0.06
p1i21	Geometry					1.71	0.07			1.58	0.07
p1i22	Geometry					1.54	0.07			1.33	0.07
p1i23	Geometry					1.65	0.08			2.15	0.09
p1i24	Algebra			1.09	0.06					1.90	0.07
p1i25	Algebra			1.26	0.06					1.01	0.05
p1i26	Algebra			1.05	0.05					-0.30	0.04
p1i27	Algebra			0.94	0.05					1.39	0.06

p1i28	Algebra			1.08	0.05					0.44	0.05
p1i29	Algebra			1.41	0.07					1.65	0.07
p1i30	Algebra			1.35	0.06					0.90	0.05
p2i1	Algebra			0.95	0.07					2.38	0.08
p2i2	Algebra			0.95	0.06					1.48	0.06
p2i3	Algebra			0.76	0.05					0.40	0.04
p2i4	Algebra			1.19	0.10					3.29	0.13
p2i5	Algebra			1.42	0.07					1.69	0.07
p2i6	Data							0.63	0.06	1.22	0.05
p2i7	Data							1.81	0.13	3.58	0.16
p2i8	Data							0.51	0.05	0.13	0.04
p2i9	Data							1.57	0.09	2.36	0.09
p2i10	Data							1.64	0.10	2.21	0.09
p2i11	Data							0.93	0.06	0.67	0.04
p2i12	Data							1.57	0.10	2.44	0.10
p2i13	Geometry					0.91	0.06			2.08	0.07
p2i14	Geometry					1.29	0.06			0.85	0.05
p2i15	Geometry					1.20	0.06			0.79	0.05
p2i16	Geometry					1.16	0.06			1.82	0.07
p2i17	Geometry			0.51	0.06	0.53	0.06			0.39	0.04
p2i18	Number	0.49	0.06	0.64	0.06					0.01	0.04
p2i19	Number	0.78	0.08	0.76	0.08					2.26	0.08
p2i20	Number	1.00	0.08	0.67	0.06					-1.00	0.05
p2i21	Number	1.00	0.16	1.56	0.19					4.85	0.28
p2i22	Number	0.76	0.07	1.02	0.07					0.64	0.05
p2i23	Number	1.27	0.07							0.11	0.05
p2i24	Data							1.14	0.07	-0.99	0.05
p2i25	Data							1.73	0.10	-0.41	0.05
p2i26	Data							1.25	0.07	0.14	0.05
p2i27	Data							1.46	0.08	-0.60	0.05
p3i1	Algebra	0.30	0.13	0.57	0.13	0.47	0.14	0.16	0.13	3.39	0.13

p3i2	Algebra	0.25	0.10	0.38	0.10	0.25	0.10	0.16	0.10	2.52	0.08
p3i3	Algebra	0.47	0.10	0.45	0.10	0.28	0.10	0.13	0.10	2.56	0.09
p3i4	Algebra	0.31	0.07	0.51	0.07	0.13	0.07	0.16	0.07	1.18	0.05
p3i5	Algebra	0.10	0.08	0.38	0.08	0.21	0.08	0.21	0.08	1.83	0.06
p3i6	Algebra	0.31	0.07	0.69	0.07	0.12	0.07	0.16	0.07	-0.82	0.05
p3i7	Algebra	0.42	0.09	0.76	0.09	0.38	0.09	0.24	0.09	2.36	0.09
p3i8	Number	0.36	0.10	0.64	0.10					2.90	0.10
p3i9	Number	0.50	0.06	0.47	0.06					1.33	0.05
p3i10	Number	0.51	0.06	0.54	0.06					0.33	0.04
p3i11	Number	0.52	0.06	0.56	0.06					0.88	0.05
p3i12	Number	0.83	0.07	0.65	0.06					1.15	0.05
p3i13	Number	0.56	0.06	0.65	0.06					0.79	0.05
p3i14	Number	1.29	0.07							1.03	0.06
p3i15	Number	1.22	0.06							0.82	0.05
p3i16	Number	0.86	0.06							1.42	0.06
p3i17	Number	0.65	0.05							0.87	0.04
p3i18	Geometry					1.47	0.11			3.42	0.14
p3i19	Geometry					0.91	0.05			1.53	0.06
p3i20	Geometry					1.11	0.08			2.58	0.09
p3i21	Geometry					0.96	0.05			0.62	0.05
p3i22	Geometry					1.17	0.06			0.83	0.05
p3i23	Geometry					1.13	0.07			2.27	0.08
p3i24	Geometry					1.12	0.06			1.22	0.06
p3i25	Geometry					0.92	0.05			1.15	0.05
p3i26	Algebra			1.62	0.08					2.36	0.09
p3i27	Algebra			0.85	0.05					0.01	0.04
p3i28	Algebra			0.83	0.05					1.22	0.05
p3i29	Algebra			0.38	0.04					-0.93	0.04

Table A16. *S-X^2 Item-level Fit from Original Blueprint based CMIRT Model*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	68.41	54	0.0895
p1i2	Number	74.62	60	0.0968
p1i3	Number	57.26	61	0.6130
p1i4	Number	91.83	62	0.0082
p1i5	Number	85.99	57	0.0078
p1i6	Number	96.99	60	0.0018
p1i7	Number	68.17	57	0.1475
p1i8	Number	71.82	57	0.0893
p1i9	Number	153.35	59	0.0001
p1i10	Number	60.52	61	0.4943
p1i11	Number	65.27	61	0.3302
p1i12	Number	91.57	60	0.0054
p1i13	Data	135.11	62	0.0001
p1i14	Data	65.41	58	0.2346
p1i15	Data	129.32	60	0.0001
p1i16	Data	105.80	61	0.0003
p1i17	Data	94.80	61	0.0036
p1i18	Data	123.64	59	0.0001
p1i19	Geometry	160.98	60	0.0001
p1i20	Geometry	89.41	61	0.0103
p1i21	Geometry	130.83	60	0.0001
p1i22	Geometry	170.68	60	0.0001
p1i23	Geometry	74.43	60	0.0994
p1i24	Algebra	80.05	60	0.0428
p1i25	Algebra	75.46	60	0.0859
p1i26	Algebra	109.34	59	0.0001
p1i27	Algebra	68.77	61	0.2305
p1i28	Algebra	133.04	59	0.0001
p1i29	Algebra	98.42	59	0.0010
p1i30	Algebra	128.44	60	0.0001
p2i1	Algebra	82.32	60	0.0295
p2i2	Algebra	86.91	61	0.0163
p2i3	Algebra	88.13	61	0.0131
p2i4	Algebra	78.35	57	0.0318
p2i5	Algebra	89.93	59	0.0058
p2i6	Data	75.61	61	0.0985
p2i7	Data	104.66	54	0.0001
p2i8	Data	97.68	62	0.0026
p2i9	Data	138.68	58	0.0001
p2i10	Data	137.39	59	0.0001
p2i11	Data	183.96	60	0.0001
p2i12	Data	131.71	58	0.0001
p2i13	Geometry	74.83	61	0.1096

p2i14	Geometry	233.35	60	0.0001
p2i15	Geometry	202.22	60	0.0001
p2i16	Geometry	113.86	60	0.0001
p2i17	Geometry	137.96	61	0.0001
p2i18	Number	110.58	60	0.0001
p2i19	Number	57.66	58	0.4890
p2i20	Number	203.99	55	0.0001
p2i21	Number	65.28	45	0.0256
p2i22	Number	105.02	58	0.0002
p2i23	Number	82.04	58	0.0206
p2i24	Data	222.38	59	0.0001
p2i25	Data	126.41	59	0.0001
p2i26	Data	154.87	61	0.0001
p2i27	Data	220.21	59	0.0001
p3i1	Algebra	88.88	55	0.0026
p3i2	Algebra	61.33	60	0.4293
p3i3	Algebra	87.72	60	0.0113
p3i4	Algebra	74.22	61	0.1190
p3i5	Algebra	56.30	61	0.6473
p3i6	Algebra	174.81	59	0.0001
p3i7	Algebra	82.52	59	0.0233
p3i8	Number	73.14	58	0.0867
p3i9	Number	61.62	61	0.4548
p3i10	Number	72.28	61	0.1527
p3i11	Number	62.03	60	0.4028
p3i12	Number	54.12	60	0.6899
p3i13	Number	61.89	60	0.4095
p3i14	Number	82.84	60	0.0270
p3i15	Number	69.70	60	0.1831
p3i16	Number	69.47	60	0.1884
p3i17	Number	57.08	62	0.6538
p3i18	Geometry	78.41	56	0.0256
p3i19	Geometry	85.62	61	0.0205
p3i20	Geometry	94.44	60	0.0030
p3i21	Geometry	146.92	61	0.0001
p3i22	Geometry	128.72	60	0.0001
p3i23	Geometry	72.56	60	0.1279
p3i24	Geometry	74.84	60	0.0938
p3i25	Geometry	107.50	61	0.0002
p3i26	Algebra	121.87	57	0.0001
p3i27	Algebra	135.24	61	0.0001
p3i28	Algebra	84.88	61	0.0233
p3i29	Algebra	222.67	61	0.0001
52 out of 86 items fitted well at the p-level of 0.01				

Table A17. *S-X^2 Item-level Fit from Augmented Blueprint based CMIRT Model*

Item	**Blueprint Standard**	X^2	**d.f.**	**Probability**
p1i1	Number	66.08	49	0.0522
p1i2	Number	79.90	59	0.0363
p1i3	Number	56.05	60	0.6214
p1i4	Number	92.66	61	0.0055
p1i5	Number	56.03	55	0.4372
p1i6	Number	136.13	60	0.0001
p1i7	Number	65.12	54	0.1426
p1i8	Number	98.60	58	0.0007
p1i9	Number	198.11	60	0.0001
p1i10	Number	72.72	62	0.1653
p1i11	Number	72.85	61	0.1421
p1i12	Number	94.40	61	0.0039
p1i13	Data	93.73	59	0.0027
p1i14	Data	62.65	56	0.2517
p1i15	Data	73.68	55	0.0470
p1i16	Data	59.83	58	0.4101
p1i17	Data	50.20	57	0.7265
p1i18	Data	50.21	52	0.5455
p1i19	Geometry	195.77	60	0.0001
p1i20	Geometry	104.47	61	0.0004
p1i21	Geometry	182.17	60	0.0001
p1i22	Geometry	220.01	60	0.0001
p1i23	Geometry	108.67	60	0.0001
p1i24	Algebra	69.68	60	0.1836
p1i25	Algebra	45.03	60	0.9249
p1i26	Algebra	75.66	59	0.0708
p1i27	Algebra	60.81	61	0.4838
p1i28	Algebra	112.52	59	0.0001
p1i29	Algebra	71.82	59	0.1218
p1i30	Algebra	93.04	59	0.0031
p2i1	Algebra	76.64	60	0.0724
p2i2	Algebra	68.14	60	0.2196
p2i3	Algebra	75.11	61	0.1055
p2i4	Algebra	67.06	56	0.1476
p2i5	Algebra	58.97	59	0.4776
p2i6	Data	107.13	61	0.0002
p2i7	Data	171.32	55	0.0001
p2i8	Data	109.62	62	0.0002
p2i9	Data	235.00	60	0.0001
p2i10	Data	230.60	60	0.0001
p2i11	Data	267.26	61	0.0001
p2i12	Data	216.08	59	0.0001
p2i13	Geometry	96.46	61	0.0026

p2i14	Geometry	278.56	60	0.0001
p2i15	Geometry	250.26	60	0.0001
p2i16	Geometry	148.50	60	0.0001
p2i17	Geometry	84.69	59	0.0158
p2i18	Number	113.06	59	0.0001
p2i19	Number	53.98	57	0.5900
p2i20	Number	190.90	51	0.0001
p2i21	Number	41.61	40	0.4020
p2i22	Number	60.60	57	0.3466
p2i23	Number	124.39	59	0.0001
p2i24	Data	201.49	60	0.0001
p2i25	Data	147.40	61	0.0001
p2i26	Data	188.69	61	0.0001
p2i27	Data	222.48	60	0.0001
p3i1	Algebra	71.50	50	0.0246
p3i2	Algebra	52.44	56	0.6111
p3i3	Algebra	70.35	55	0.0794
p3i4	Algebra	59.54	58	0.4205
p3i5	Algebra	50.66	58	0.7426
p3i6	Algebra	152.38	52	0.0001
p3i7	Algebra	53.43	53	0.4587
p3i8	Number	64.04	57	0.2429
p3i9	Number	54.92	59	0.6274
p3i10	Number	71.77	59	0.1227
p3i11	Number	59.09	58	0.4367
p3i12	Number	50.28	57	0.7239
p3i13	Number	49.46	58	0.7805
p3i14	Number	94.74	60	0.0028
p3i15	Number	76.00	60	0.0795
p3i16	Number	66.92	60	0.2512
p3i17	Number	61.24	62	0.5045
p3i18	Geometry	99.93	56	0.0003
p3i19	Geometry	105.55	61	0.0003
p3i20	Geometry	117.49	60	0.0001
p3i21	Geometry	174.84	61	0.0001
p3i22	Geometry	160.12	61	0.0001
p3i23	Geometry	97.23	60	0.0017
p3i24	Geometry	101.97	61	0.0008
p3i25	Geometry	136.66	61	0.0001
p3i26	Algebra	85.35	57	0.0089
p3i27	Algebra	110.36	60	0.0001
p3i28	Algebra	66.66	61	0.2881
p3i29	Algebra	224.96	61	0.0001
50 out of 86 items fitted well at the p-level of 0.001				

Table A18. *Guess and Slip Parameter Estimates from DINA Model*

Item	Original Blueprint-based Standards	Guess	SE(Guess)	Slip	SE(Slip)
p1i1	Number	0.932	0.007	0.014	0.003
p1i2	Number	0.785	0.011	0.031	0.005
p1i3	Number	0.438	0.013	0.289	0.012
p1i4	Number	0.694	0.012	0.456	0.013
p1i5	Number	0.661	0.012	0.031	0.005
p1i6	Number	0.374	0.013	0.145	0.009
p1i7	Number	0.906	0.008	0.015	0.003
p1i8	Number	0.775	0.011	0.027	0.004
p1i9	Number	**0.288**	0.013	0.180	0.010
p1i10	Number	0.395	0.013	0.262	0.012
p1i11	Number	0.729	0.012	0.062	0.006
p1i12	Number	**0.245**	0.012	0.479	0.013
p1i13	Data	0.396	0.012	0.239	0.012
p1i14	Data	0.879	0.008	0.017	0.004
p1i15	Data	0.697	0.012	0.053	0.006
p1i16	Data	0.520	0.013	0.143	0.010
p1i17	Data	0.682	0.012	0.052	0.006
p1i18	Data	0.730	0.011	0.021	0.004
p1i19	Geometry	**0.356**	0.014	0.128	0.009
p1i20	Geometry	0.500	0.014	0.098	0.008
p1i21	Geometry	0.426	0.014	0.052	0.006
p1i22	Geometry	0.406	0.014	0.072	0.007
p1i23	Geometry	0.569	0.014	0.037	0.005
p1i24	Algebra	0.646	0.014	0.064	0.006
p1i25	Algebra	0.402	0.014	0.133	0.009
p1i26	Algebra	**0.182**	0.011	0.374	0.012
p1i27	Algebra	0.574	0.014	0.111	0.008
p1i28	Algebra	**0.305**	0.014	0.219	0.010
p1i29	Algebra	0.502	0.014	0.062	0.006
p1i30	Algebra	**0.341**	0.014	0.127	0.009
p2i1	Algebra	0.781	0.012	0.049	0.005
p2i2	Algebra	0.593	0.014	0.104	0.008
p2i3	Algebra	0.400	0.014	0.283	0.011
p2i4	Algebra	0.857	0.010	0.014	0.003
p2i5	Algebra	0.519	0.014	0.068	0.007
p2i6	Data	0.658	0.013	0.143	0.009
p2i7	Data	0.849	0.010	0.006	0.002
p2i8	Data	0.414	0.014	0.349	0.013
p2i9	Data	0.713	0.012	0.025	0.004
p2i10	Data	0.673	0.013	0.025	0.004
p2i11	Data	0.425	0.014	0.150	0.010
p2i12	Data	0.725	0.012	0.019	0.004
p2i13	Geometry	0.724	0.012	0.054	0.006

p2i14	Geometry	**0.353**	0.013	0.129	0.009
p2i15	Geometry	**0.354**	0.013	0.141	0.009
p2i16	Geometry	0.602	0.014	0.053	0.006
p2i17	Geometry	0.369	0.013	0.225	0.011
p2i18	Number	**0.297**	0.012	0.284	0.012
p2i19	Number	0.708	0.012	0.022	0.004
p2i20	Number	**0.106**	0.008	0.416	0.013
p2i21	Number	0.895	0.008	0.002	0.001
p2i22	Number	**0.339**	0.012	0.125	0.009
p2i23	Number	**0.259**	0.012	0.236	0.011
p2i24	Data	**0.179**	0.011	0.539	0.013
p2i25	Data	**0.222**	0.012	0.347	0.013
p2i26	Data	**0.315**	0.013	0.254	0.012
p2i27	Data	**0.192**	0.011	0.395	0.013
p3i1	Algebra	0.895	0.008	0.017	0.004
p3i2	Algebra	0.848	0.009	0.036	0.005
p3i3	Algebra	0.821	0.010	0.033	0.005
p3i4	Algebra	0.587	0.012	0.102	0.009
p3i5	Algebra	0.764	0.011	0.075	0.008
p3i6	Algebra	**0.169**	0.009	0.434	0.014
p3i7	Algebra	0.720	0.011	0.022	0.004
p3i8	Number	0.887	0.008	0.033	0.005
p3i9	Number	0.642	0.012	0.124	0.009
p3i10	Number	0.372	0.013	0.230	0.011
p3i11	Number	0.500	0.013	0.145	0.010
p3i12	Number	0.497	0.013	0.094	0.008
p3i13	Number	0.465	0.013	0.155	0.010
p3i14	Number	0.454	0.014	0.113	0.009
p3i15	Number	0.432	0.014	0.150	0.010
p3i16	Number	0.653	0.013	0.121	0.009
p3i17	Number	0.559	0.014	0.195	0.011
p3i18	Geometry	0.830	0.010	0.013	0.003
p3i19	Geometry	0.626	0.013	0.104	0.008
p3i20	Geometry	0.769	0.012	0.028	0.004
p3i21	Geometry	0.396	0.014	0.207	0.010
p3i22	Geometry	0.386	0.014	0.151	0.009
p3i23	Geometry	0.717	0.012	0.046	0.005
p3i24	Geometry	0.497	0.014	0.119	0.008
p3i25	Geometry	0.518	0.014	0.128	0.009
p3i26	Algebra	0.606	0.014	0.026	0.004
p3i27	Algebra	**0.284**	0.013	0.342	0.012
p3i28	Algebra	0.557	0.014	0.132	0.009
p3i29	Algebra	**0.216**	0.012	0.658	0.012

Table A19. *MLTM-D Item Difficulty Estimates on Components*

Item	Blueprint Standard	b1	SE(b1)	b2	SE(b2)	b3	SE(b3)	b4	SE(b4)
p1i1	Number	-6.08	1.17	-3.45	0.18				
p1i2	Number	-3.29	0.20	-2.84	0.16				
p1i3	Number	-0.91	0.15	-1.86	0.25				
p1i4	Number	-0.66	0.05	-6.50	.				
p1i5	Number	-4.02	0.41	-1.82	0.09				
p1i6	Number	-0.67	0.05						
p1i7	Number	-4.31	0.31	-3.64	0.24				
p1i8	Number	-2.48	0.07						
p1i9	Number	-0.35	0.05						
p1i10	Number	-0.40	0.05						
p1i11	Number	-2.08	0.07						
p1i12	Number	0.57	0.05						
p1i13	Data	-0.81	0.13	-2.38	0.40			-2.86	0.77
p1i14	Data	-3.60	0.20					-3.64	0.31
p1i15	Data	-3.05	0.22	-2.32	0.15			-3.76	0.53
p1i16	Data	-2.26	0.22	-1.67	0.15			-2.89	0.38
p1i17	Data	-3.25	0.27	-2.42	0.16			-2.88	0.26
p1i18	Data	-3.48	0.26	-2.49	0.15			-3.71	0.44
p1i19	Geometry					-0.64	0.04		
p1i20	Geometry					-1.04	0.04		
p1i21	Geometry					-0.99	0.04		
p1i22	Geometry					-0.90	0.04		
p1i23	Geometry					-1.39	0.04		
p1i24	Algebra			-1.65	0.05				
p1i25	Algebra			-0.80	0.04				
p1i26	Algebra			0.31	0.04				
p1i27	Algebra			-1.28	0.05				
p1i28	Algebra			-0.36	0.04				
p1i29	Algebra			-1.26	0.05				
p1i30	Algebra			-0.69	0.04				
p2i1	Algebra			-2.19	0.06				
p2i2	Algebra			-1.35	0.05				
p2i3	Algebra			-0.38	0.04				
p2i4	Algebra			-2.81	0.08				
p2i5	Algebra			-1.29	0.05				
p2i6	Data							-1.22	0.05
p2i7	Data							-2.54	0.08
p2i8	Data							-0.15	0.04
p2i9	Data							-1.80	0.06
p2i10	Data							-1.66	0.06
p2i11	Data							-0.64	0.04
p2i12	Data							-1.87	0.06
p2i13	Geometry					-1.79	0.05		

p2i14	Geometry					-0.63	0.04		
p2i15	Geometry					-0.60	0.04		
p2i16	Geometry					-1.43	0.04		
p2i17	Geometry			-1.03	0.16	-1.33	0.19		
p2i18	Number	-1.63	0.26	-0.56	0.14				
p2i19	Number	-3.43	0.25	-2.26	0.12				
p2i20	Number	-0.45	0.17	-0.13	0.13				
p2i21	Number	-6.05	1.21	-3.15	0.14				
p2i22	Number	-2.50	0.25	-0.78	0.09				
p2i23	Number	-0.13	0.05						
p2i24	Data							0.78	0.04
p2i25	Data							0.23	0.04
p2i26	Data							-0.17	0.04
p2i27	Data							0.40	0.04
p3i1	Algebra	-6.08	1.37	-3.52	0.22	-3.43	0.23	-6.50	.
p3i2	Algebra	-4.27	0.41	-3.47	0.28	-3.45	0.27	-3.97	0.49
p3i3	Algebra	-3.89	0.31	-3.46	0.29	-3.25	0.25	-4.10	0.53
p3i4	Algebra	-2.69	0.25	-1.97	0.17	-3.37	0.50	-3.19	0.45
p3i5	Algebra	-3.30	0.27	-2.94	0.28	-3.25	0.28	-3.73	0.69
p3i6	Algebra	-1.69	0.41	0.39	0.13	-5.76	4.35	-5.74	3.55
p3i7	Algebra	-5.17	1.08	-2.47	0.16	-2.72	0.19	-4.11	0.63
p3i8	Number	-4.12	0.28	-3.25	0.18				
p3i9	Number	-2.25	0.16	-2.16	0.16				
p3i10	Number	-1.62	0.22	-1.01	0.15				
p3i11	Number	-1.93	0.18	-1.62	0.14				
p3i12	Number	-2.17	0.17	-1.62	0.12				
p3i13	Number	-2.06	0.19	-1.33	0.12				
p3i14	Number	-0.98	0.05						
p3i15	Number	-0.81	0.05						
p3i16	Number	-1.56	0.06						
p3i17	Number	-1.03	0.05						
p3i18	Geometry					-2.43	0.06		
p3i19	Geometry					-1.33	0.04		
p3i20	Geometry					-2.07	0.05		
p3i21	Geometry					-0.52	0.04		
p3i22	Geometry					-0.64	0.04		
p3i23	Geometry					-1.80	0.05		
p3i24	Geometry					-0.97	0.04		
p3i25	Geometry					-1.00	0.04		
p3i26	Algebra			-1.68	0.05				
p3i27	Algebra			0.02	0.04				
p3i28	Algebra			-1.17	0.05				
p3i29	Algebra			1.05	0.05				

Table A20. X^2 *Item-level Fit from MLTM-D*

Item	Blueprint Standard	X^2	d.f.	Probability
p1i1	Number	2.033	4	0.730
p1i2	Number	1.302	8	0.996
p1i3	Number	59.155	13	0.000
p1i4	Number	279.020	12	0.000
p1i5	Number	22.686	11	0.020
p1i6	Number	33.316	12	0.001
p1i7	Number	1.217	4	0.875
p1i8	Number	13.822	7	0.054
p1i9	Number	35.289	13	0.001
p1i10	Number	9.067	13	0.768
p1i11	Number	4.977	9	0.836
p1i12	Number	34.762	11	0.000
p1i13	Data	31.620	13	0.003
p1i14	Data	3.751	5	0.586
p1i15	Data	16.228	11	0.133
p1i16	Data	13.026	14	0.524
p1i17	Data	8.175	12	0.771
p1i18	Data	21.427	10	0.018
p1i19	Geometry	12.386	14	0.575
p1i20	Geometry	4.502	14	0.992
p1i21	Geometry	30.188	14	0.007
p1i22	Geometry	18.427	14	0.188
p1i23	Geometry	15.368	13	0.285
p1i24	Algebra	2.195	11	0.998
p1i25	Algebra	12.307	14	0.582
p1i26	Algebra	10.177	13	0.679
p1i27	Algebra	6.113	12	0.910
p1i28	Algebra	36.359	14	0.001
p1i29	Algebra	15.943	12	0.194
p1i30	Algebra	28.364	14	0.013
p2i1	Algebra	6.554	9	0.683
p2i2	Algebra	7.005	12	0.857
p2i3	Algebra	26.158	14	0.025
p2i4	Algebra	4.572	6	0.600
p2i5	Algebra	20.480	12	0.059
p2i6	Data	21.697	13	0.060
p2i7	Data	14.778	7	0.039
p2i8	Data	61.748	13	0.000
p2i9	Data	23.951	11	0.013
p2i10	Data	30.864	11	0.001
p2i11	Data	11.458	14	0.650
p2i12	Data	24.267	10	0.007
p2i13	Geometry	5.290	11	0.916

p2i14	Geometry	12.160	14	0.593
p2i15	Geometry	6.190	14	0.961
p2i16	Geometry	4.912	13	0.977
p2i17	Geometry	55.404	14	0.000
p2i18	Number	30.456	13	0.004
p2i19	Number	9.366	10	0.498
p2i20	Number	46.539	12	0.000
p2i21	Number	10.149	5	0.071
p2i22	Number	36.157	14	0.001
p2i23	Number	35.530	13	0.001
p2i24	Data	71.853	10	0.000
p2i25	Data	53.346	13	0.000
p2i26	Data	29.289	13	0.006
p2i27	Data	64.186	10	0.000
p3i1	Algebra	12.291	5	0.031
p3i2	Algebra	3.551	7	0.830
p3i3	Algebra	12.037	8	0.150
p3i4	Algebra	9.079	13	0.767
p3i5	Algebra	5.515	10	0.854
p3i6	Algebra	29.643	12	0.003
p3i7	Algebra	9.657	11	0.561
p3i8	Number	4.015	6	0.675
p3i9	Number	9.041	12	0.699
p3i10	Number	23.681	14	0.050
p3i11	Number	12.062	14	0.601
p3i12	Number	9.041	14	0.828
p3i13	Number	6.257	14	0.960
p3i14	Number	33.611	13	0.001
p3i15	Number	26.872	13	0.013
p3i16	Number	6.262	11	0.855
p3i17	Number	6.426	12	0.893
p3i18	Geometry	9.874	8	0.274
p3i19	Geometry	9.167	13	0.760
p3i20	Geometry	5.969	10	0.818
p3i21	Geometry	20.828	14	0.106
p3i22	Geometry	16.011	14	0.313
p3i23	Geometry	6.973	11	0.801
p3i24	Geometry	12.820	14	0.541
p3i25	Geometry	10.858	14	0.697
p3i26	Algebra	25.769	11	0.007
p3i27	Algebra	60.005	13	0.000
p3i28	Algebra	14.263	13	0.356
p3i29	Algebra	258.222	11	0.000
61 out of 86 items fitted well at the p-level of 0.01				

REFERENCES

Bolt, D. M., & Lall, V. F. (2003). Estimation of compensatory and noncompensatory multidimensional item response models using Markov chain Monte Carlo. *Applied Psychological Measurement, 29*, 395- 414.

Bock, R. D., & Aiken, M. (1981). Marginal maximum likelihood estimation of item parameters: application of an EM algorithm, *Psychometrika*, *46*, 443-445.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1988). Full information item factor analysis. *Applied Psychological Measurement, 12*, 261–280.

de la Torre, J., & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika, 69*, 333-353.

de la Torre, J., & Douglas, J. (2005). Modeling multiple strategies in cognitive diagnosis. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada.

DiBello, L. V., Stout, W. F., & Roussos, L. A. (1995). Unified cognitive/psychometric diagnostic assessement likelihoon-based classification techniques. Chapter 15 in Nichols, P. D., Chipman, S. F. and Brennan, R. L. (Eds.) (1995). *Cognitively diagnostic assessment.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Embretson, S.E. (2000). Multidimensional measurement from dynamic tests: Abstract reasoning under stress. *Multivariate Behavioral Research 35*, 505-543.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179-197.

Embretson, S. E. (1984). A general multicomponent latent trait model for response processes. *Psychometrika, 49,* 175-186.

Embretson, S. E. (1997). Multicomponent latent trait models. In W. J. van der Linden & R. K. Hambleton (Eds), *Handbook of modern item response theory* (pp. 305-322). New York: Springer.

Embretson, S. E. (2000). Multidimensional measurement from dynamic tests: Abstract reasoning under stress. *Multivariate Behavioral Research, 35*, 505-543.

Embretson, S. E., & McCollam, K. M. (2000). A multicomponent Rasch model for measuring covert processes: Application to lifespan ability changes. In M. Wilson & G. Engelhard(Eds.), *Objective Measurement: Theory into practice* (Vol. 5, pp. 203-218) Upper Saddle River, NJ: Ablex.

Embretson, S. E., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Embretson. S. E., & Yang. X. (2013). A multicomponent latent trait model for diagnosis. *Psychometrika, 78*(1), 14-36.

Gierl, M. J. & Zhou, J. (2008). Computer adaptive-attribute testing: A new approach to cognitive diagnostic assessment. *Journal of Pyschology, 216* (1), 29-39.

Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement, 26*, 333-352.

Hartz, S. M. (2002). *A Bayesian framework for the Unified Model for assessing cognitive abilities: Blending theory with practicality*. Unpublished doctoral dissertation, University of Illinois.

Hartz, S., Roussos, L., & Stout, W. (2002). Skills diagnosis: Theory and practice [User manual for Arpeggio software]. Princeton, NJ: Educational Testing Service.

Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log linear models with latent variables. *Psychometrika, 74*, 191-210.

Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*, 258-272.

Orlando, M., & Thissen, D. (2000). Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement, 24*, 50-64.

Pellegrino, J. W. (2002). Understanding how students learn and inferring what they know: Implications for the design of curriculum, instruction, and assessment. In M. J. Smith (Ed.), *NSF K-12 Mathematics and science curriculum and implementation centers conference proceedings* (pp. 76-92). Washington, DC: National Science Foundation and American Geological Institute.

Reckase, M. D., & McKinley, R. L. (1982). *The feasibility of a multidimensional latent trait model.* Paper presented at the meeting of the American Psychological Association, Washington.

Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement, 15*, 361-373.

Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications.* New York, NY: Guilford Press.

Schilling, S., & Bock, R. D. (2005). High-dimensional maximum marginal likelihood item factor analysis by adaptive quadrature. *Psychometrika, 70,* 533-555.

Stegelmann, W. (1983). Expanding the Rasch model to a general model having more than one dimension. *Psychometrika, 48*, 259-267.

Stone, C. A., & Zhang, B. (2003). Assessing goodness of fit of item response theory models: a comparison of traditional and alternative procedures. *Journal of Educational Measurement, 40*, 331-352.

Stout, W. (2007). Skills diagnosis using IRT-based continous latent trait models. *Journal of Educational Measurement, 44*, 313-324.

Thissen, D. & Cai, L. (2011). IRTPRO (Beta version) [Computer software]. Chicago, IL: Scientific Software International.

von Davier, M. (2005). *A general diagnostic model applied to language testing data* (ETS Research Rep. No. RR-05-16). Princeton, NJ: ETS.

Von Davier, M., & Yamamoto, K. (2007). Mixture distribution Rasch models and hybrid Rasch models. In M. Davier., & C. H. Carstensen (Eds.). *Multivariate and mixture distribution. Rasch models*. New York, NY: Springer.

Wainer, H., Vevea, J.L., Camacho, F., Reeve, B., Rosa, K., Nelson, L., Swygert, K., & Thissen, D. (2001). Augmented scores—"borrowing strength" to compute scores based on small numbers of items. In D. Thissen & H. Wainer (Eds), *Test Scoring* (pp. 343-387). Mahwah, NJ: Lawrence Erlbaum.

Whitely, S. E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika, 45*, 479-494.

Wilson, D. T, Wood, R., & Gibbons, R. (2003). TESTFACT 4.0: Test scoring, item statistics, and item factor analysis [Computer software]. Chicago, IL: Scientific Software International.

Yang, X., & Embretson, S. E. (2007). Construct validity and cognitive diagnostic assessment. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 119 -145). New York: Cambridge University Press